INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

Neutrophil Gelatinase-associated Lipocalin: A Protein Involved in Immune Defense Against Microbial Pathogens

By
David Henry Goetz

A dissertation submitted in partial fulfillment of the requirements of the degree of

Doctor of Philosophy

University of Washington

2002

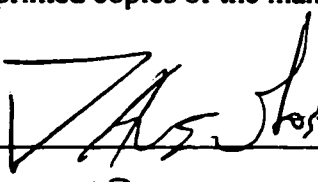
Program Authorized to Offer Degree: Molecular and Cellular Biology Program

UMI Number: 3053503

UMI Microform 3053503

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

In presenting this thesis in partial fulfillment of the requirements for the Doctoral degree at the University of Washington, I agree that the Library shall make its copies freely available for inspection. I further agree that extensive copying of the dissertation is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for copying or reproduction of this dissertation may be referred to ProQuest Information and Learning, 300 North Zeeb Road, P.O. Box 1346, Ann Arbor, MI 48106-1346, to whom the author has granted "the right to reproduce and sell (a) copies of the manuscript in microform and/or (b) printed copies of the manuscript made from microform."

Signature 

Date June 13, 2002

University of Washington
Graduate School

This is to certify that I have examined this copy of a doctoral dissertation by

David Henry Goetz

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Chair of Supervisory Committee:

Roland K. Strong

Reading Committee:

Roland K. Strong

Jefferson Foote

David Teller

Date: June 13, 2002

University of Washington

Abstract

Structural and Biochemical Studies of Neutrophil Gelatinase-associated Lipocalin: A Protein Involved in Immune Defense Against Microbial Pathogens

by David Henry Goetz

Chair Person of the Supervisory Committee: Affiliate Assistant Professor Roland K. Strong, Department of Immunology, Associate Member Fred Hutchinson Cancer Research Center

Human neutrophil gelatinase-associated lipocalin was discovered a decade ago as a protein secreted in response to inflammatory stimuli. NGAL was originally described as a component of a disulphide linked heterodimer with gelatinase-B isolated from specific granules secreted by neutrophils. NGAL expression patterns suggested that it plays a role in many processes including cell-growth, apoptosis, cell-differentiation, neoplasia, and defense against bacterial infection. Sequence analysis suggested that NGAL was a member of the lipocalin structural family. The lipocalins are characterized by an eight-stranded β-barrel, which forms a cup-shaped -generally lipophilic- binding site referred to as the calyx. NGAL was first proposed to function by binding small bacterial hydrophobic chemotactic peptides. However, biochemical data in support of this hypothesis has been difficult to obtain. The aim of my thesis project has been to use structural and biochemical techniques to determine the function of NGAL. My crystal structures of baculovirus-expressed monomer and disulphide-linked homodimer NGAL suggested that NGAL does not in fact bind hydrophobic chemotactic peptides. Subsequently, I serendipitously discovered that NGAL co-purifies with a bacterial chromophore when expressed heterologously in *E. coli* XL1-Blue. The remainder of my thesis project has been to characterize this chromophore and study its possible relationship to NGAL's *in vivo* function. I identified the bacterial chromophore as

enterobactin and showed that the ligand tightly associates with the NGAL calyx. Enterobactin is a member of a class of iron-binding small molecules known as siderophores that are secreted by microorganisms for the purpose of scavenging iron. My subsequent crystal structure of the NGAL:enterobactin complex revealed that the binding mode exhibits some unique properties. First, NGAL appears to have a binding site that is not only capable of tightly associating with enterobactin but also other larger siderophores. Second, NGAL appears to primarily recognize the catecholate moieties of enterobactin through unique hybrid cation-π/ionic interactions. These data suggest that NGAL may be part of a general mechanism of the innate immune system for inhibiting bacterial catecholate-siderophore-mediated iron-scavenging in a mammalian host.

TABLE OF CONTENTS

List of Figures

List of Tables

Acknowledgments

The author wishes to thank the following people: My parents above all, those zany folks over in the Parkhurst Lab, Susan Parkhurst for great advice, and most of all my friends who have made graduate school not only tolerable, but a great deal of fun as well.

Chapter 1: Introduction to Neutrophil Gelatinase-associated Lipocalin

Neutrophils

The focus of this thesis is on structural and biochemical characterization of the human protein neutrophil gelatinase-associated lipocalin (NGAL). NGAL was first described as a protein secreted by polynuclear effector cells of the innate immune system known as neutrophils. Neutrophils are the most common leukocyte in the blood and bone marrow comprising up to 48% of all white cells in the blood of normal individuals[1]. After the physical barriers of the skin and mucosal layers, neutrophils are the first line of defense against infection by microorganisms. Neutrophils function by chemotaxing towards sites of infection and inflammation; directly killing invading microorganism by phagocytosis, formation of reactive oxidants, and release of bactericidal and bacteristatic peptides; and secreting a number of inflammatory mediators to modulate the activity of other components of the immune system. Inability to maintain normal levels of circulating active neutrophils leads to increased susceptibility to infection by gram-negative bacteria. At levels below 1000 circulating neutrophils per µl of blood, there is an inverse relationship between neutrophil count and risk of bacterial infection[2].

Neutrophils are derived from the myeloid lineage, and so neutrophil precursors are found in the bone marrow[3]. Neutrophils mature from myelocytes to metamyelocytes, to immature, or band, neutrophils and finally to segmented (mature) neutrophils[4]. Mature neutrophils leave the marrow, circulate through the blood, and periodically loosely attach to the vascular endothelium. Once attached, neutrophils either, detach and return to the circulating blood, or extravasate across the endothelial layer into the surrounding tissues. This decision is mediated by a number of both soluble and membrane bound inflammatory signaling molecules. Neutrophils spend an average of 6 to 10 hours in the blood before entering extravascular tissues or body cavities[5].

One hallmark of infections by gram-negative bacteria such as *E. coli* is a brief decrease in the level of circulating neutrophils, a condition referred to as neutropenia. This brief neutropenia is usually followed by a marked increase in the number of

circulating neutrophils (neutrophilia). In healthy patients, neutrophilia is accompanied by an increase in the proliferation and maturation of neutrophils in the bone marrow throughout the course of infection. Resolution of the infection is associated with a return to normal levels of neutrophil production. In cases of severe infection and/or patients with pre-existing bone marrow deficiencies, the kinetics of the proliferation and maturation of neutrophils may be too slow keep up with the death of neutrophils occurring extravascularly. This situation quickly leads to a reduction in the pool of proliferating neutrophil precursors and ultimately a deficiency of phagocytic cells at the sites of infection. The prognosis under such a condition is generally poor.

In order to fend off pathogens, neutrophils must be able to migrate to the site of invasion. Neutrophil migration first begins with loose attachment to vascular endothelial cells. This attachment is loose enough to allow the flow of blood through the vessel to roll the neutrophil over the surface of the endothelial cells. This attachment is mediated by the binding of P type selectins on the surface of neutrophils and L type selectins on endothelial cells to their corresponding carbohydrate ligands on the opposite cell[6]. After the initial attachment, neutrophils flatten out, and attach more strongly to the endothelial cells[7]. This step is mediated by the binding of activated $ß_2$ integrins on the surface of the neutrophil to intercellular adhesion molecules (ICAMs) on the endothelial cells[8]. The activation of $ß_2$ integrin is mediated by pro-inflammatory signals originating from the endothelial cells, such as platelet-activating factor (PAF), interleukin-8 (IL-8), histamine, and leukotriene-B_4 (LTB4) [7]. Following attachment, neutrophils migrate between the endothelial cells into the extravascular tissues through interaction with platelet-endothelial cell adhesion molecule 1[8].

Neutrophils must next chemotax towards the specific site of infection after extravasation. This process is mediated by interaction with a number of signaling molecules known as chemotaxins[9]. These chemotaxins include LTB4, the complement cleavage fragment C5a, and the tri-peptide N-formyl-methionine-leucine-phenylalanine (fMLF) [10]. Human neutrophils possess approximately 50,000 fMLF receptors on their cell surface[11]. Binding of fMLF results in elongation and polarization of the neutrophil[7].

Polarized neutrophils have a large ruffled leading edge with a high concentration of fMLF receptors and a narrow trailing edge with a low concentration of receptors. Following polarization, neutrophils crawl over the surface substrate, to which they are loosely attached, along a chemotactic gradient.

When a neutrophil contacts an invading pathogen, the pathogen is ingested and the killing cascade begins. During phagocytosis, neutrophil granules fuse to the membrane of the phagocytic vesicle as well as the cytoplasmic membrane. As a result of fusion, granule contents are released into the phagosome and the extracellular matrix. Neutrophil granules are large proteinaceous bodies first identified by microscopy, which contain an array of proteins involved in host defense against microbes. There are three distinct types of granules: azurophil or primary granules, secondary or specific granules, and tertiary granules[12]. All three types release their contents into the phagosome, but only specific granules are released extracellularly[13]. The identification of granules as primary, secondary, and tertiary is somewhat arbitrary as there is overlap in both protein content and kinetics of release. However, primary granules are considered those that contain myeloperoxidase. Secondary granules are identified by the specific marker lactoferrin, a protein that functions to bind iron with high affinity[14]. Tertiary granules are poorly characterized, but are noted for their gelatinase content[15].

Granular fusion and release of contents into the phagosome initiates a process called the respiratory burst or oxidative killing[16]. This process leads to the formation of reactive oxidants toxic to the invading pathogen. The initial step is the oxidation of reduced nicotinamide adenine dinucleotide phosphate (NADPH) to NADP by a membrane bound enzyme in the phagosome[17]. The product of this reaction is the superoxide anion $O_2\cdot^-$. $O_2\cdot^-$ can form hydrogen peroxide (H_2O_2) either spontaneously or through the activity of the enzyme superoxide dismutase[18,19]. Myeloperoxidase then converts H_2O_2 into HOCl; H_2O_2 also spontaneously converts into the hydroxyl radical $OH\cdot$ in the presence of iron[20]. H_2O_2, $OH\cdot$, and HOCl are all potent bactericidal chemicals. In addition to proteins involved in oxidative bacterial killing, primary granules also

contain small cationic polypeptides called defensins[21], which form pores in the bacterial cell membrane, and lysozyme, which degrades bacterial cell walls[22].

Isolation Identification and Primary Structure of NGAL

NGAL was first discovered over a decade ago as a disulphide linked component of a 125 kilodalton (kDa) gelatinase-B containing complex isolated from human neutrophils[23]. Treibel *et al.* purified gelatinase-B from neutrophils on gelatin sepharose and eluted NGAL with a reducing buffer. The protein was identified by tryptic digestion and N-terminal sequencing. Subsequently, NGAL was identified as a protein released from the granules of phorbol myristate acetate (PMA) stimulated neutrophils[24]. NGAL was described as a 25 kDA protein exocytosed by stimulated neutrophils that cross reacts with polyclonal antibodies raised against gelatinase-B. PMA selectively induces the release of specific neutrophil granules[25]. NGAL was later purified as a 45 kDa disulphide linked homodimer isolated from specific granules[26]. Xu *et al.* named the protein they isolated human neutrohil lipocalin (HNL) and reported that stimulated neutrophils exocytose approximately 290 fg of NGAL/HNL per cell or 51% of total NGAL stores. Thus NGAL was identified as a principal component of specific granules where it exists in three forms: a monomer, a disulphide linked homodimer, and a disulphide linked heterodimer with gelatinase-B. Approximately 10% of NGAL in neutrophils exists as the monomer form, 50% as the homodimer and 50% in complex with gelatinase-B. Of the NGAL in complex with gelatinase-B, 50% is actually in a ternary complex with tissue inihibitor of protease 1[27]. Epithelial cells do not express gelatinase-B, and thus there is no heterodimer in NGAL secreted from epithelia.

NGAL is a 178 amino acid protein with one N-linked glycosylation site at N65. Treatment of NGAL with glycolytic enzymes results in a 20 kDa protein, which agrees with the calculated mass[24]. The sequence of NGAL shows a high degree of homology with the mouse protein 24p3 (62% identical) and the rat protein α_2-microglobulin related protein (63.5% identical). The rat and mouse homologues had previously been shown to

be members of a structural family known as lipocalins[28]. Thus NGAL was named neutrophil gelatinase-associated lipocalin.

NGAL Homologues

α_2-microglobulin related protein was first reported in rats as a fortuitously isolated cDNA with % 52 identity to α_2-microglobulin/ major urinary protein[29]. Subsequently, the messenger RNA was found to be 12-fold overexpressed in *neu* induced mammary tumors compared to *ras* or chemically induced tumors and the protein was named *neu*-related lipocalin (NRL) [30]. The *neu* oncogene is unique in that it is a much more potent inducer of mammary tumors than *ras*[31]; Although many genes were overexpressed in both *ras* and *neu* induced tumors, only NRL showed a significantly higher expression level by differential hybridization in *neu* induced tumors over *ras* induced tumors. These data suggest a role for NRL, and thus NGAL, in the regulation of oncogenesis and cell proliferation.

The mouse homologue of NGAL, 24p3,was first described as a gene induced to 14-fold overexpression by the SV40 virus in mouse kidney cells arrested in G_0[32]. G_0 arrested kidney cells infected with SV40 undergo a mitotic host reaction and begin to divide rapidly. These data suggest that 24p3, and thus NGAL, may be involved in the regulation of cell proliferation. Indeed, 24p3 has been found to be expressed in response to a number of different mitogenic signals including the glucocortocoid dexamethasone and retinoic acid[33]. Other researchers described the same protein as a superinducible protein isolated from 3T3 cells stimulated with fibroblast growth factor and named it SIP24[34]. Subsequently, the mouse protein was also named utrerocalin because it is highly expressed in the luminal fluid and epithelia of the uterus during parturition[35,36]. 24p3 is induced to secretion in mouse macrophages by exposure to inflammatory signals such as lipopolysaccharide[37]. Macrophages and neutrophils play complementary roles in extravascular tissue during invasion by microorganisms. The data by Meheus *et al.* suggest that 24p3 and NGAL share functional as well as sequence homology in that they

both are implicated in host defense against bacterial pathogens. More recently, ectopic recombinant 24p3 has been shown to induce apoptosis in several T-cell lines under IL-3 withdrawal. [38]

Tissue Distribution of NGAL 24p3 and NRL

The precise localization of mRNA expression, and protein secretion often provides valuable clues about the function of a gene. It has been proposed that NGAL is a scavenger of bacterial products at sites of inflammation[39]. This hypothesis suggests that NGAL may act by binding to and sequestering inflammatory signals of bacterial origin from the rest of the immune system; thus, down-regulating the immune response. In order for NGAL to function by binding to inflammatory signals of bacterial origin, it must be secreted from cells exposed to microorganisms. Similarly, if NGAL plays a role in cell proliferation, or carcinogenesis, it should be expressed in cells undergoing such a transformation.

Originally, NGAL mRNA expression was tested in fifty physiologically normal human adult tissues by dot-blot hybridization[40]. The results showed a high level of gene expression in bone marrow, but not in circulating blood cells. These data confirm earlier observations that neutrophils express and package NGAL into granules during development in the bone marrow at the myelocyte and metamyelocyte stage[15]. Thus NGAL is not expressed by mature neutrophils, rather it is expressed during development, stored in granules, and secreted by mature neutrophils in response to inflammatory signals. Significant mRNA expression levels were also found in the uterus, prostate, salivary gland, stomach, appendix, colon, trachea, and lung. Furthermore, NGAL expression was induced in bronchial epithelia and alveoli during inflammation (Cowland *et al.* manuscript submitted). These data support the idea that NGAL is involved in host defense against microorganisms; since all tissues that show NGAL expression are either directly exposed to microorganisms, or excrete to organs which are exposed.

Enhanced expression of NGAL is detected in the colon epithelia of patients with inflammatory bowel disease and neoplasia[39]. Similarly, although normal ovarian cells

express little NGAL, ovarian cancer cells exhibit high NGAL expression[41]. NGAL is also constitutively expressed and secreted by A549 cells, a cell line derived from a human alveolar cell carcinoma (Cowland *et al.* manuscript submitted). These results indicate that NGAL may play a role in carcinogenesis and the regulation of cell proliferation. With few exceptions, the mRNA expression patterns of 24p3 and NRL mirror those of NGAL. Furthermore, some of the expression patterns observed for 24p3 and NRL in mice and rats have not been looked for in human tissues and may exist there as well. This is particularly the case for developing embryos. In rats, NRL expression is seen in a variety of developing tissues including: growth plate cartilage, prehypertrophic chondrocytes, skeletal muscle fibers, and the myocardium of the developing heart[42]. Analysis of NGAL protein levels largely agree with the mRNA expression level studies[43]. NGAL was reported in the tissues of the lung, stomach, small intestine, pancreas, kidney, prostate gland, and thymus. Friedl *et al.* also showed that NGAL and NRL protein expression were very similar with the exception of the pancreas and kidney. The presence of NGAL in rat neutrophils has not yet been investigated.

Lipocalin Structure and Function

NGAL and its homologues 24p3 and NRL are members of the structural family of proteins known as lipocalins[28]. The lipocalins comprise an extremely large family of proteins exhibiting a great array of functional variation. Nevertheless, certain common functional, structural, and sequence motifs are conserved. Lipocalins are small, single domain, secreted proteins which bind small, generally hydrophobic molecules[44]. The overall fold of a lipocalin is an eight-stranded continuously hydrogen-bonded ß-barrel, the interior of which forms a cup shaped binding pocket referred to as a calyx(Figure 1) [45]. The strands of the barrel are all linked by +1 connections, the simplest ß-sheet topology. The loops that connect the ß-strands are typical ß-hairpin loops, except the loop between strands 1 and 2, designated loop A. Typically, loop A is a large Ω loop that tends to fold back over the binding pocket and form a lid over the interior of the calyx. Between the last two strands there is a conserved α-helix.

The so-called kernel lipocalins, of which NGAL is a member, share three characteristic conserved sequence motifs[46]. Sequences within the three motifs are almost invariant throughout the kernel lipocalins; however the sequence of NGAL does in fact differ quite substantially in these regions. The first motif (FQDNQFQGKWYVVGL) consists of an N-terminal 3_{10}-helix which packs against a conserved helix in strand 1. The second motif (LVRVVSTNYNQHAMVFFK) consists of strands six and seven and the short linking loop between them. The final motif (YFKITLYGRTK) consists of the C-terminus of strand eight and the loop that connects to the conserved α-helix (LH) as well as residues in the N-terminal half of the helix. For readers wishing a more comprehensive review of lipocalin struture and function please refer to the following reviews[46,47].

Proposed Function of NGAL and Supporting Biochemical Data

NGAL's classification as a lipocalin, its exocytosis by neutrophils in response to inflammatory stimuli, and its expression in tissues exposed to microorganisms suggested that NGAL might bind to a small hydrophobic inflammatory signal either from bacteria or secreted by the host in response to bacterial infection. A number of possible ligands have been proposed in the literature. These include PAF, LTB4, IL-8, fatty acids, and fMLF[48].

Prior to the first identification of NGAL[26,49], the protein was actually partially purified and characterized as a 24 kDa component of specific granules exocytosed by neutrophils in response to PMA or fMLF that binds to fMLF[50]; however, the protein was not actually identified as NGAL until much later[51]. Both groups used a heterobifunctional compound sulfosuccinimidyl 2-(p-azidosalicylamido)ethyl-1-3'-dithiopropionate (SASD) (Figure 2) to radiolabel fMLF and look for proteins isolated from neutrophils which bound to the labeled compound fMLF-SASD. Sengelov *et al.* reported an equilibrium dissociation constant (K_d) of 1.4×10^{-7} M for the labeled NGAL:fMLF-SASD complex determined by measurement of free and bound label at equilibrium, and a K_d of 5×10^{-5} M for the unlabeled NGAL:fMLF complex determined by competition of labeled ligand binding with unlabeled ligand. Several factors including

the greater affinity of NGAL for fMLF-SASD over fMLF indicated that NGAL was actually interacting with the SASD label. Furthermore, Sengelov *et al.* showed that NGAL exhibited no preference for fMLF over the non-formylated peptide Met-Leu-Phe, which does not have any affect on neutrophil chemotaxis or inflammatory response. Later, researchers reported similarly weak binding for 24p3 to fMLF by fluorescence quenching with a reported K_d of $2x10^{-6}$ M[52]. However, the tryptophan fluorescence quenching experiments used a large concentration of tryptophanamide as the solvent for the stock fMLF solution. It is quite likely that the measured interaction between NGAL and ligand was between NGAL and tryptophanamide since later fluorescence experiments using other solvents failed to show any concentration dependant fluorescence quenching. The most reliable estimates to date indicate an extremely weak affinity for NGAL to fMLF. These estimates are based on corrected retention times of the proposed ligands through an immobilized NGAL column matrix. The technique relies on the fact that the interaction is so weak that k_{off} is large enough and k_{on} small enough to allow rapid movement of the ligand through the matrix. The best estimate of the affinity of NGAL for fMLF is a K_d of $1.1x10^{-3}$ M[48]. This affinity is extremely low, and suggests that perhaps NGAL is not involved in binding to fMLF. Bratt *et al.* reported similarly weak affinities for the other proposed NGAL ligands PAF (K_d=$3.9x10^{-5}$M) and LTB4 (K_d= $2.3x10^{-4}$M).

X-ray Crystallography

X-ray crystallography is a powerful tool for determining detailed structural information of biological macromolecules at atomic resolution. The following sections provide an overview of the theory and practice behind macromolecular X-ray crystallography. This is not a complete description of all aspects of modern crystallographic techniques. For a more thorough review of modern crystallographic theory and methods, the reader is invited to consult the following sources[53,54].

When a molecule is irradiated with high-energy photons (X-rays), the electrons in the molecule are accelerated by collisions with the incident radiation. The electrons then

re-emit this radiation. Crystallography primarily deals with the special case of elastic collisions, which result in an emitted photon with a uniform change of phase and an identical wavelength with respect to the incident photon.

To describe the scattering from any object, it is necessary to describe the electron density (usually electrons per $Å^3$) as a function of position ($\rho(r)$). If we consider incident photons of wavelength λ, then the path difference between X-rays scattered by electrons at the origin and at any given point r is given by equation (1) $r \bullet (k - k_0)$ where k_0 is the wave vector describing the incident radiation, k describes the scattered radiation and $|k_0| = \frac{1}{\lambda}$. The phase difference between scattered and incident radiation is given by equation (2) $2\pi i r \bullet (k - k_0)$. Thus, the scattered wave of a volume dr is given by equation (3) $\rho(r) e^{2\pi i r \bullet s} dr$ where s is the scattering vector k-k_0. Integration over the entire volume of the molecule provides an equation to describe the scattering of the whole molecule. This is called the structure factor equation $F(s) = \int \rho(r) e^{2\pi i r \bullet s} dr$.

$F(s)$ is a Fourier transform of the electron density distribution $\rho(r)$. One extremely useful property of any Fourier transform is that its inverse is also true; *viz.*, $\rho(r) = A \int F(s) e^{-2\pi i r \bullet s} ds$ where A is a constant to put electron density on an absolute scale. Therefore, if we can experimentally measure all photons scattered by an irradiated object ($F(s)$), then calculation of the inverse Fourier transform of $F(s)$ would yield the electron-density distribution of the object, and thus its atomic structure. Unfortunately, $F(s) = |F(s)| e^{i\phi}$ where $|F(s)|$ is the amplitude of the structure factor and ϕ is the phase. While $|F(s)|$ can be readily measured, the phase information is lost during that measurement. The inability to measure this phase is called the phase problem.

In a crystal with unit cell axes **a, b, and c**, electron density ($\rho(r)$) is a three dimensional periodic function of position with translational periods **a**, **b**, and **c**. The crystal lattice acts as a three dimensional diffraction grating in which scattered X-rays are subject to constructive and destructive interference. In order for measurable diffraction to occur as a result of constructive interference, the path difference between photons striking adjacent periodic elements in the crystal must be an integral number of the

wavelength λ. Bragg's law, equation (4) $\sin(\theta) = \frac{n\lambda}{2d}$ where 2(θ) is the angle through which the photons are scattered, describes the values which satisfy this condition simultaneously in all three dimensions. The structure factor equation for a crystal then becomes $F(h,k,l) = \sum_{n=1}^{N} f_n(s) e^{2\pi i(hx_n + ky_n + lz_n)}$ where f_n is the atomic structure factor and h, k, l are the reciprocal lattice constants.

The value *d* in equation (4) is a measure of the detail of the information content from a given measurement. For large scattering angles, d is correspondingly small and the level of detail measured is high. The minimum value of d used to determine a particular structure is termed the resolution of that structure. As the resolution of the observed data increases, so does the accuracy of the resultant structure. The resolution of the observed data is a function of the internal order of the crystal, the strength and quality of the incident radiation, and the size of the crystal.

Macromolecular crystal structures are not perfect for several reasons: the atomic model generally only describes a static structure with isotropic thermal motion, low resolution characteristics of the crystal such as bulk solvent are difficult to model accurately, and phases cannot be measured only estimated. For these reasons, there will always be a measurable difference between the observed amplitudes of the structure factors F_{obs} (F_o) and the stucture factor amplitudes calculated from the atomic model by an inverse Fourier transform F_{calc} (F_c). This difference is called the crystallographic R-factor and is given in equation (5) $R = \frac{\sum |F_{obs} - F_{calc}|}{\sum |F_{obs}|}$. High R-values denote a large difference between F_{obs} and F_{calc}, while low R-values denote a small difference. In practice, R-values below about 0.25 are considered acceptable for medium resolution macromolecular protein crystal structures.

The difference between F_{obs} and F_{calc} can be exploited for computational refinement of the coordinates and thermal properties of the atomic model. Many different refinement strategies have been reported. The strategies in common use today, however, can be generalized by an expression of the form $\Phi = \varphi_1 + \varphi_2 \ldots$ where the terms φ_i correspond to different variables to be minimized. For example, φ_1 could be of the form

of equation (6) $Q = \sum_{hkl} w(hkl)\left\{|F_{obs}(hkl)| - k|F_{calc}(hkl)|\right\}^2$ where w is a weighting factor for the reliability of $F_{obs}(hkl)$ and k is a scale factor. Similarly ϕ_2 could be an equation describing the known range of values for bond-lengths and bond-angles within amino acids determined from previously solved high-resolution crystal structures. The final equation Φ is then minimized by least squares or maximum likelihood methods.

One interesting feature of the refinement of atomic positions is that large values of $|F_{obs} - F_{calc}|$ that arise from local errors in atomic positions can be minimized by shifts distant from the actual site of error. Therefore, while $|F_{obs} - F_{calc}|$ and the R-factor get smaller, the model in fact gets worse. This problem has necessitated the use of an R_{free} value[55]. The R_{free} is calculated exactly like the R-factor except that a subset of around one-thousand F_{obs} that were measured and not used in refinement are utilized for the calculation. Thus, R_{free} represents an independent validation of the accuracy of the atomic model. To minimize confusion, the R-factor that is calculated from the remaining F_{obs}, which are refined against, is often referred to as the R_{work}.

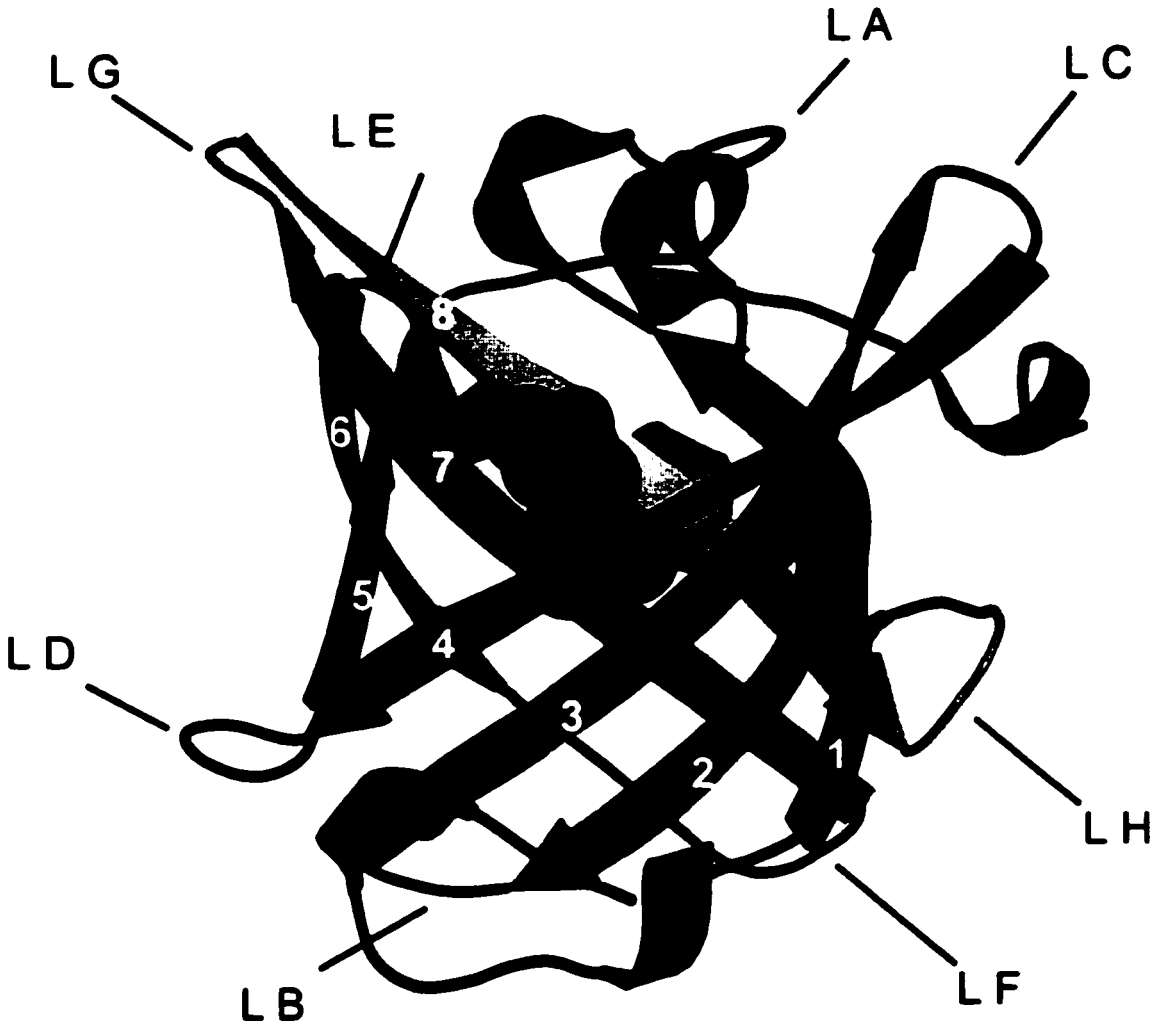


Figure 1: Ribbon Diagram of the Kernel Lipocalin Mouse Major Urinary Protein (MUP). Strands are denoted by thick sheets capped with arrows, helices are denoted by thin sheets, coils by thin tubes, and the ligand hydroxy-methyl-heptanone is depicted in cpk colored by atom type: red for oxygen and black for carbon. The strands that form the ß-barrel are numbered 1 – 8, and the loops between them are labeled LA - LG.

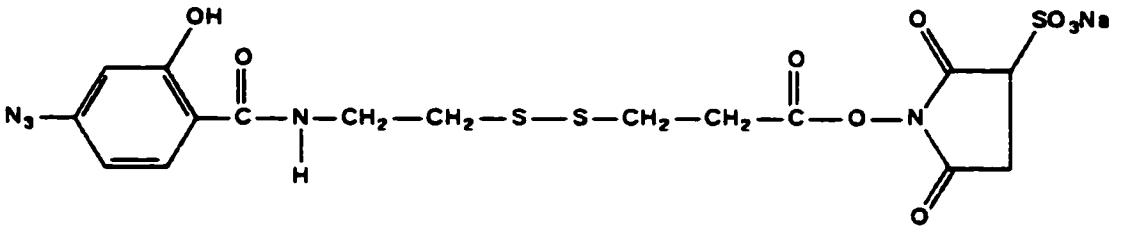


Figure 2: Structure of the Heterobifunctional Crosslinking Agent SASD. Note the *P*-Azidosalicylamido functional group on the left.

Chapter 2: The Structure of Neutrophil Gelatinase-associated Lipocalin

Structure Determination of Monomer and Dimer NGAL

NGAL was expressed, purified and crystallized as previously described[56]. Briefly, NGAL was expressed and purified from Sf9 insect cells using the BaculoGold baculovirus based expression system (Pharmingen). Protein was expressed and purified by antibody affinity chromatography in the laboratory of Niels Borregaard, followed by gel filtration to resolve dimer and monomer forms in our lab. The monomer form of NGAL was treated with iodoacetamide to carboxymethylate free sulfhydryl groups and maintain the monomeric state of the protein. The use of traditional reducing agents such as dithiothreitol (DTT) instead of iodoacetamide was precluded by the fact that such reducing agents are rather easily oxidized and degrade during long-term storage.

NGAL homodimer crystals were grown by vapor diffusion at 22°. The reservoir solution was 20% polyethylene glycol 8000, 100mM sodium acetate pH 4.5, and 15% glycerol. The protein solution was 20mg/ml NGAL homodimer in a 25mM neutral pH buffer. The drop was setup as a one to one mixture of protein and reservoir solution, and crystals appeared within one week. NGAL homodimer crystallized in spacegroup $P4_12_12$ with unit cell edges a=b=54.29Å, and c=122.0Å and half of a dimer in the asymmetric unit. NGAL monomer crystals were grown by vapor diffusion from a solution of 10mg/ml protein in 25mM neutral pH buffer against a reservoir of 2M ammonium sulfate, and 10% glycerol. Monomer crystal quality was improved by the addition of 5-10mM dithiothreitol in the protein drop.

The dimer structure was determined using estimated phases from single isomorphous replacement and anomalous scattering. The isomorphous derivative was obtained by a 24 h soak in a solution of 20% polyethylene glycol 8000, 100mM sodium acetate pH 4.5, 15% glycerol, and 1mM di-*m*-iodobis(ethylenediamine)-diplatinum(II) nitrate (PIP). Data were collected at 100 K on a Rigaku R-axis 4c image plate, and processed with DENZO and SCALEPACK[57]. Heavy atom sites were found by manual inspection of difference Pattersons. Heavy atom parameters were refined in SHARP[58].

The resulting phases were extended from 3.9 Å to 2.5 Å by solvent flattening in SHARP via the solomon[59] routine. The initial electron density map showed obvious solvent/macromolecule boundaries and secondary structural elements. However, the precise strand connectivity and placement was very ambiguous. Furthermore, many of the loop regions contained no interpretable electron density. The final model was obtained by iterative model building using XtalView[60] and refinement with CNS[61].

The structure of the monomer was determined by molecular replacement using a partially refined structure of the dimer as a search model in the program EPMR[62]. Two solutions were found (designated monmer A and monomer B) yielding a final R-factor after rigid body minimization of 35.8%. Non-crystallographic restraints were not applied during refinement because R_{free} values were consistently lower in the absence of restraints. The monomer structures were completed by iterative cycles of manual building and computational refinement. Intermediate structures were manually compared against sigmaa[63] weighted $2F_o$-F_c, F_o-F_c, and composite simulated annealing omit maps. The geometry of the models was evaluated with the procheck[64] package. Residues 1- 3 in the dimer and monomer B and residues 1-4 in monomer A are disordered. The following additional residues were modeled as alanines due to the poor quality of the local electron density: T-4, E-61, K-62, R-72, K-73, K-74, K-98 in the dimer structure; and T-4 in monomer B. Data collection and refinement parameters for the monomer and dimer structures of baculovirus expressed NGAL can be found in Table 1.

Table 1: Data Collection and Refinement Statistics for Baculoviral NGAL

Data Set	Dimer	PIP	Monomer
Data Collection			
Resolution (Å)	2.4	2.85	2.5
Unique Reflections	6872	4509	25147
Redundancy	6.43	6.39	4.2
Completeness (%)	88.1 (38.7)	90.2 (49.7)	87.9 (60.0)
I/σI	20.4 (5.2)	17.1 (2.2)	43.8 (7.6)
R_{sym} (%)	5.2 (10.9)	5.6 (10.9)	5.2 (25.5)
Isomorphous Difference (%)		18	
Phasing power (centrics)		2.28	

Table 1 Continued

Phasing power (acentrics)	2.73	
Figure of merit (to 3.9 Å ;%)	64.6	
Refinement		
Resolution (Å)	20-2.4	20-2.6
Protein atoms	1392	21861
Solvent atoms	74	133
Carbohydrate atoms	39	42
Ligand atoms	12	10
R_{work} (%)	21.6	28.1
R_{free} (%)	28.5	29.3
Geometry (rmsd from ideal)		
Bond length (Å)	0.008	0.008
Bond angles (deg)	1.4	1.7
Ramachandran analysis		
Residues in most favored regions (%)	77	80.7
Residues in disallowed regions	0	0

Numbers in parentheses indicate values for the highest resolution shell

Overall Structure of Neutrophil Gelatinase-associated Lipocalin

The crystal structure of dimeric human NGAL , crystallized at low ionic strength in an acidic pH buffer, was determined at 2.4 Å resolution. The crystal structure of the monomer form of NGAL was determined at 2.6 Å resolution at neutral pH in a high ionic strength buffer. There is half a dimer in the asymmetric unit of the dimer crystals, and two monomers in the asymmetric unit of the monomer crystals, yielding three independent views of the NGAL structure. In spite of the large differences in salt concentration and pH, the three NGAL structures are essentially identical. The overall root mean squared deviation (rmsd) between the Cα's of the three structures ranges from 0.24 Å between monomer A and B to 0.71 Å between monomer A and the dimer (Figure 3). NGAL exhibits a typical lipocalin fold, with the conserved eight-stranded ß-barrel, a short 3_{10}-helix (residues 24-28) that packs against a conserved tryptophan (residue 31), and a four-turn α-helix (residues 145-160). A ninth, three-residue ß-strand (residues 165-167) packs against strand 1. The N-terminus of NGAL, however, is more ordered that

most lipocalin structures and folds back against the side of the barrel passing over the sixth, seventh, and eighth ß-strands (Figure 3).

The largest difference in the three NGAL structures lies at the N-termini, where the first two ordered residues (4 and 5) of each independent structure have different conformations, affected by different crystal packing environments. The largest backbone difference in the remainder of the structure is at loop E, which moves by approximately 3 Å at P101 from the dimer to the monomer structures. P101 lies at a crystal contact in the dimer structure and comes within 4 Å of the Cγ of E147 of a neighboring molecule, which may explain the large difference in backbone position. The next largest difference is at loop C, which moves by 1.9 Å at K74 between the dimer and monomers. Within the calyx, L70, W79, and R81 are the only residues that show any significant conformational change (Figure 4). L70 adopts three different rotomers. W79 is disordered and modeled as alanine in the dimer structure, but not in the monomer structures. R81 either points towards the bottom of the calyx in the dimer structure, or upward and towards the center of the calyx in the monomer structures.

A search of the structures deposited in the protein databank via DALI[65] reveals that NGAL is most closely related structurally to epididymal retinoic acid-binding protein (ERBP) and mouse major urinary protein (MUP), and somewhat less closely related to ß-lactoglobulin (ß-LG) and nitrophorin (NP). The Cα rmsd's between NGAL and the related structures are 1.9Å, 2Å, 2.5Å, and 3.4Å respectively. The largest differences between NGAL and the other lipocalins lie at loops A and C. Loop A, rather than folding back over the open end of the calyx as in ERBP and MUP, packs against the last strand, which leads to a much wider calyx (Figures 5 and 6). Similarly, Loop C in ERBP folds back over the mouth of the calyx, but does not in NGAL (Figure 6).

In all three structures of NGAL there is a clear indication of the disulphide bond between C76 and C175 that is conserved amongst all known lipocalins. The remaining cysteine C87 provides the thiol group responsible for the formation of homodimers and heterodimers with gelatinase-B. The structure of the half dimer suggests that there may not be any functional significance of the high molecular weight disulphide-linked forms of NGAL. Formation of the disulphide bond does not occlude or affect the structure of

the calyx. C87 lies in a hairpin turn between strands 4 and 5 and it is likely that this allows a great deal of flexibility between the disulphide-linked proteins. There is no other interaction between the two monomers of the dimer other than the disulphide bond. The free cysteine in gelatinase-B (C449), to which NGAL disulphide-links, lies in what is thought to be a long flexible loop between the matrixin and homopexin domains. Thus, it is likely that the disulphide bond between NGAL and gelatinase-B has little affect on the structure of either molecule. This, however, does not preclude the hypothesis that *in vivo* NGAL homodimer acts as a cross-linking agent.

There is one ordered N-linked carbohydrate residue visible at N65 in the dimer structure, with two ordered carbohydrate residues visible in monomer A, and one in monomer B. There appears to be little interaction between the protein residues and the N-linked carbohydrates. The carbohydrate substituents in all three structures lie on the outside of the barrel and point away from the protein towards solvent. In both crystal forms, however, the carbohydrates do point somewhat obliquely towards the calyx of neighboring molecules. Gel-filtration experiments, however, did not demonstrate any affinity between NGAL and several tested glycoproteins such as lectin and fetuin, nor do NGAL monomers have any detectable affinity for each other.

Comparison of Crystal and NMR structures

The structure of NGAL was also solved in solution by nuclear magnetic resonance spectroscopy (NMR) [66]. Although both the NMR and crystal structures of NGAL reveal the same basic fold, there are a number of key differences. As a whole, the NMR structure differs little from the crystal structures although there is an overall rmsd of 4 Å on all common Cα's between the regularized mean NMR structure and the crystal structures (Figure 7). The N-terminal arm, up to residue 17 is the most divergent part of the structures, extending away from the remainder of the protein in the NMR structure. Excluding this arm from the comparison yields an rmsd of 3 Å between the regularized mean NMR structure and the crystal structures. Removal of the sections of the NMR model lacking secondary structure, leads to an overall rmsd of 1.4 Å between the crystal

structures and the NMR model. The average rmsd between the 20 convergent NMR structures (2.2 Å on all backbone atoms) is larger than the deviations between the three independent NGAL crystal structures by more than three-fold.

The NMR results suggest that the α-helix represents a rigid structural element flexibly linked to the rest of the protein; and indeed, the α-helix lies in a considerably different position as compared to the crystal structure (Figure 7). However, the crystal structures all show identical conformations for the α-helix (Figure 3); furthermore, the thermal parameters of the helix do not differ from those of the rest of the structure (Table 1). Finally, there is better agreement between the position of the α-helix in the NGAL crystal structures and α-helices of the related lipocalins MUP and ERBP (Figures 5 and 6) than between the position of the α-helix in the NMR and crystal structures. These differences are unlikely to reflect differences due to solution *versus* crystal environments since very different crystallization conditions and crystal packing environments yield essentially identical structures.

Loop C in the regularized NMR structure lies in an even more open configuration than in the NGAL crystal structure with the Cα of K74 in the dimer structure 8.9 Å closer to the center of the calyx than the NMR model. This indicates that perhaps the calyx is capable of opening even wider than the crystal structures suggest (Figure 7).

Molecules in the Calyx of the NGAL Crystal Structures

Late in the refinement process, unidentified Fourier objects (UFO's) were found in the NGAL calyx of all three structures that did not correspond to protein, ordered solvent, or carbohydrate (Figure 8). In the dimer structure, *n*-capric acid (NCA) a fatty acid with a nine carbon long tail was built into the electron density. Mass spectroscopy measurements to confirm the identification of the copurified ligand have been inconclusive. The electron density within the calyx, and especially near the NCA did not appear to change regardless of whether the proposed ligand $fMLF^{51}$ was either soaked into the crystal or co-crystallized. Although the NCA fits into the electron density (Figure 8), the fit is far from ideal, and it is likely that the molecule corresponding to the UFO is

not actually a fatty acid. However the position of the putative fatty acid does provide useful information with regard to the identity of the true NGAL ligand.

The carboxylate of the NCA hydrogen bonds to the side chains of Y52, K134, and Y138 and lies in a pocket lined by the side chains of T54, Y56, R81, F123, and T136. This pocket lies deep in the NGAL calyx (Figure 9). This pocket appears to be well suited for binding to carboxylate functional groups and bears slight resemblance to the carboxylate binding pocket found in fatty acid binding proteins[67]. However, as well as the carboxylate functional group of the fatty acid fits into the "carboxylate pocket" of the calyx of NGAL, the NCA molecule fails to fill the rest of the calyx (Figures 9 and 10). Finally, the positively charged lysine and arginine residues lining the calyx fail to complement the hydrophobic nature of the fatty acid tail.

Late in the refinement of the monomer crystal form of NGAL, another UFO was discovered. This UFO was centered much higher up towards the mouth of the calyx, and appeared to interact with the Nζ's of K125, and K134 as well as the guanidinium group of R81, and the Nε of W79. The shape of the UFO, the high concentration of ammonium sulfate in that crystallization condition, and the charge of the interacting arginine and lysine side chains led to the identification of the monomer UFO as a sulfate ion. The presence of the ordered sulfate ion appears to lead to a number of side chain movements in the calyx that occur for both monomers in the asymmetric unit (Figure 4). First, W79 becomes ordered and orients towards the middle of the calyx such that the Nε points toward the sulfate. Second, the side chain of R81 flips out of a hydrogen bond to the side chain of S68 on the inner wall of the ß-barrel, towards the center of the calyx to interact with the sulfate ion.

Implications of Structure on Proposed Function

The binding pockets of most lipocalins are small and hydrophobic, suitable for binding small, hydrophobic ligands (Figure 10). The binding pocket of NGAL, however, is unusually large and positively charged. This suggests that the actual ligand for NGAL is in fact much larger than the usual lipocalin ligands such as retinoic acid, and negatively

charged. The shape and charge characteristics that would complement the NGAL calyx rule out any of the proposed ligands for NGAL.

The NCA and sulfate molecules found in the NGAL calyx provide clues to the identity of the true NGAL ligand. First the location and charge of the sulfate ion, suggests that the calyx of NGAL is uniquely suited to interact with negatively charged ligands. Second, the position of the carboxylate functional group of the fatty acid suggests that this pocket within the calyx is likely to bind carboxylate-like moieties. Finally, the large areas within the calyx that both the sulfate and the fatty acid fail to fill suggest that the ligand for NGAL is much larger than previously thought.

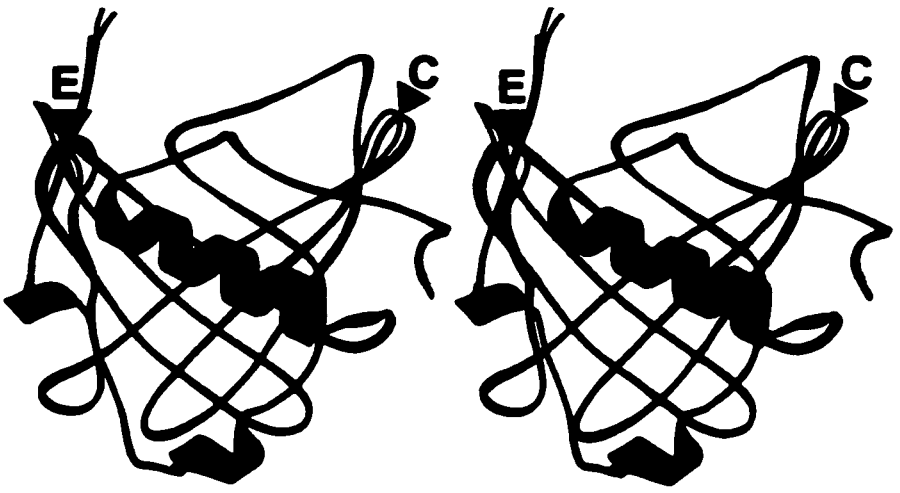


Figure 3: The Crystal Structure of NGAL. Stereo view of the two monomer and one-half dimer structures superposed and colored by chain: monomer A is blue, monomer B is red, and the half dimer is purple. Loops C and E are labeled with a black triangle. The N-terminus is colored yellow, and the C-terminus is colored green.

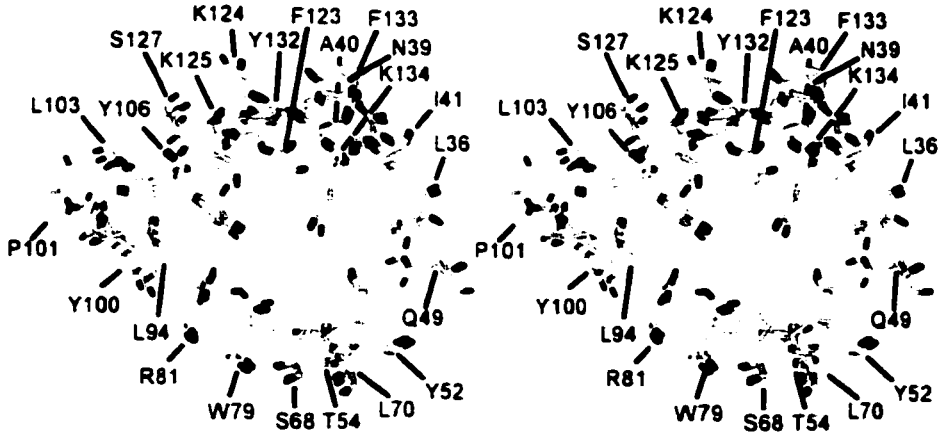


Figure 4. Residues of the NGAL Calyx. Stereo diagram of all three structures superposed with respect to their Cα positions. Atoms are colored according to type: oxygen atoms are red, nitrogen atoms are blue, and carbon atoms are gray.

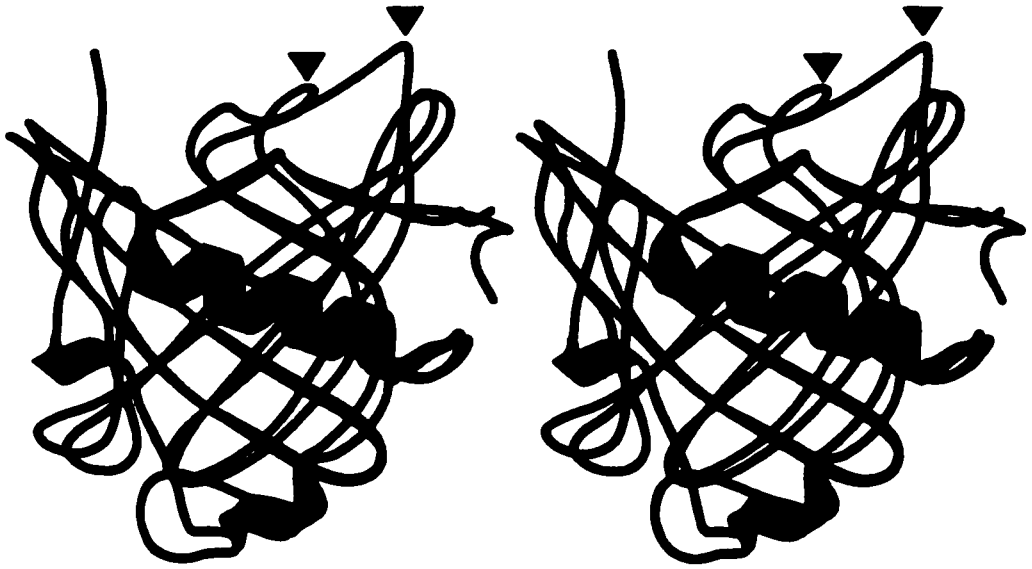


Figure 5: Stereo View of MUP (green) Superposed onto NGAL Dimer (purple). Loop A is labeled with a black triangle for both structures.

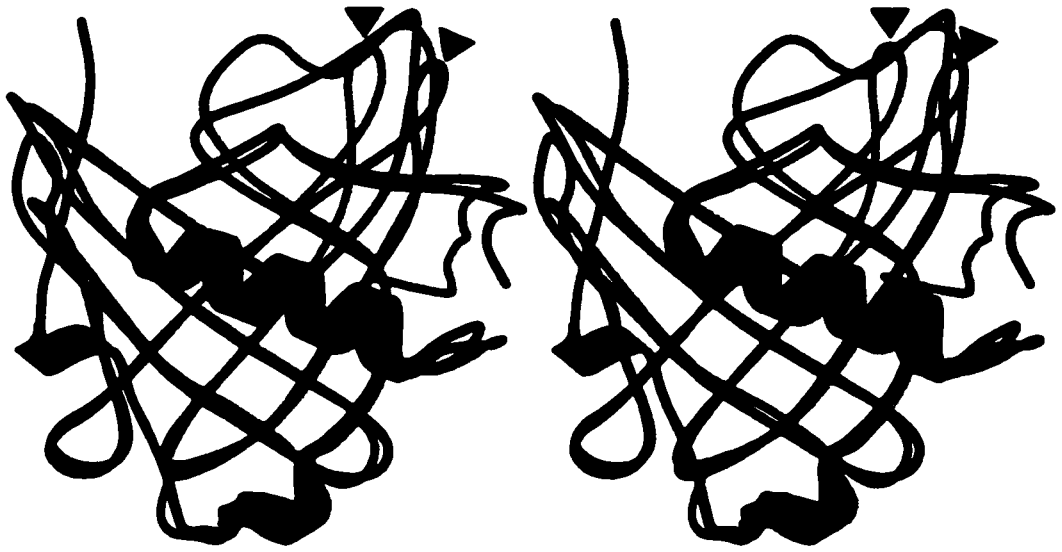


Figure 6: Stereo View of Epididymal Retinoic-acid Binding Protein (ERBP orange) Superposed onto NGAL Dimer (purple). Loop C is labeled with a black triangle.

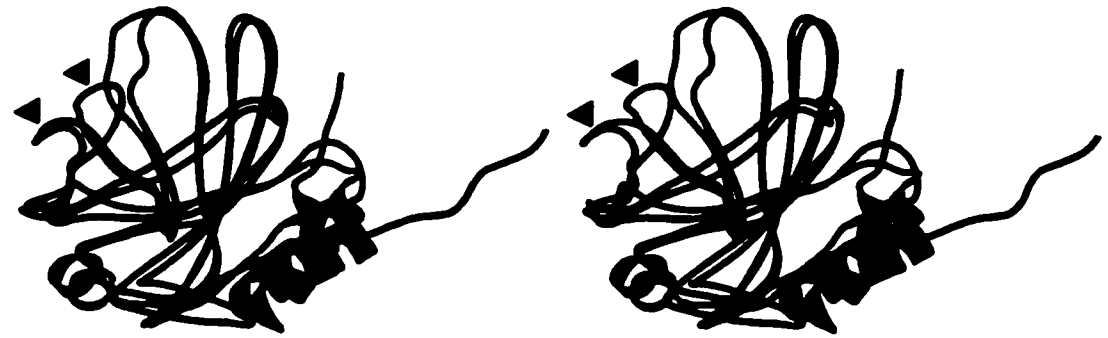


Figure 7: Stereo View of the Dimer Crystal Structure (purple) and the Regularized Mean NMR Structure (brown) Superposed. The N and C termini are colored as in Figure 3. Loop C is labeled with a black triangle.

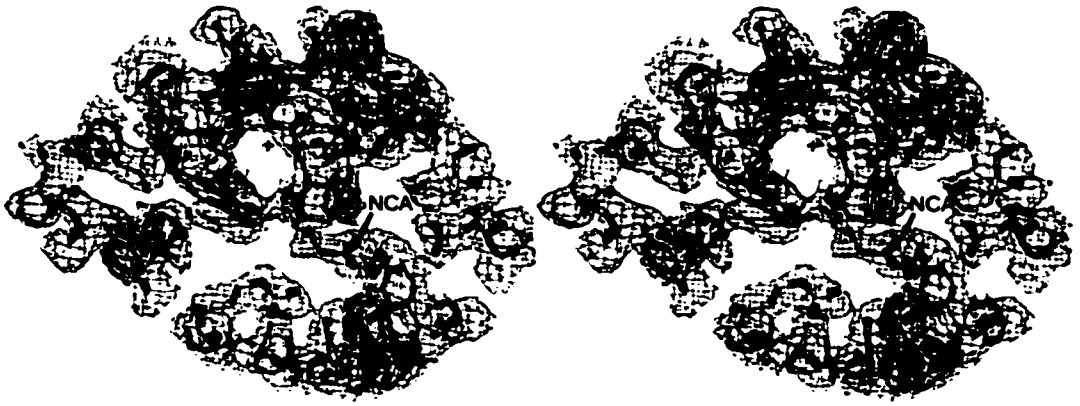


Figure 8: Stereo View of $2F_{obs}$-F_{calc} Electron Density in the Binding Pocket of the Dimer. The fatty acid is labeled (NCA). This figure is in the same orientation as Figure 4.

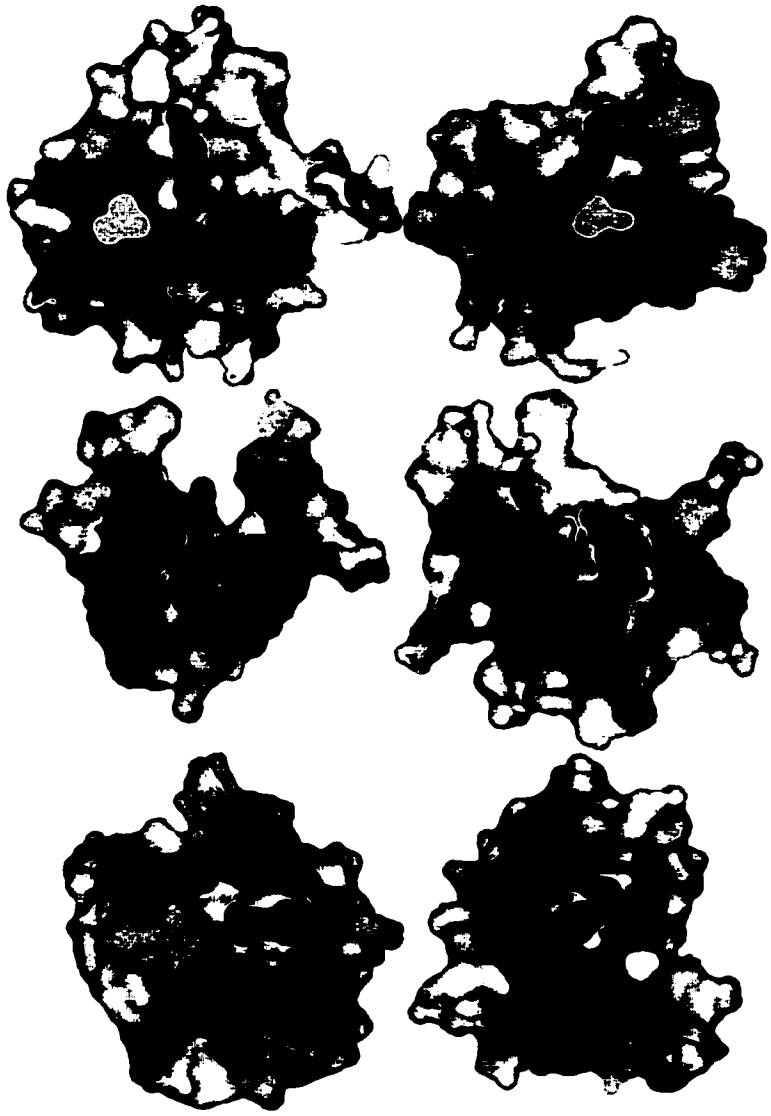


Figure 9: Orthogonal 4 Å Thick Slices Through Molecular Surfaces of NGAL (red), ERBP (orange), and MUP (green). The ligands in each case have been colored blue.

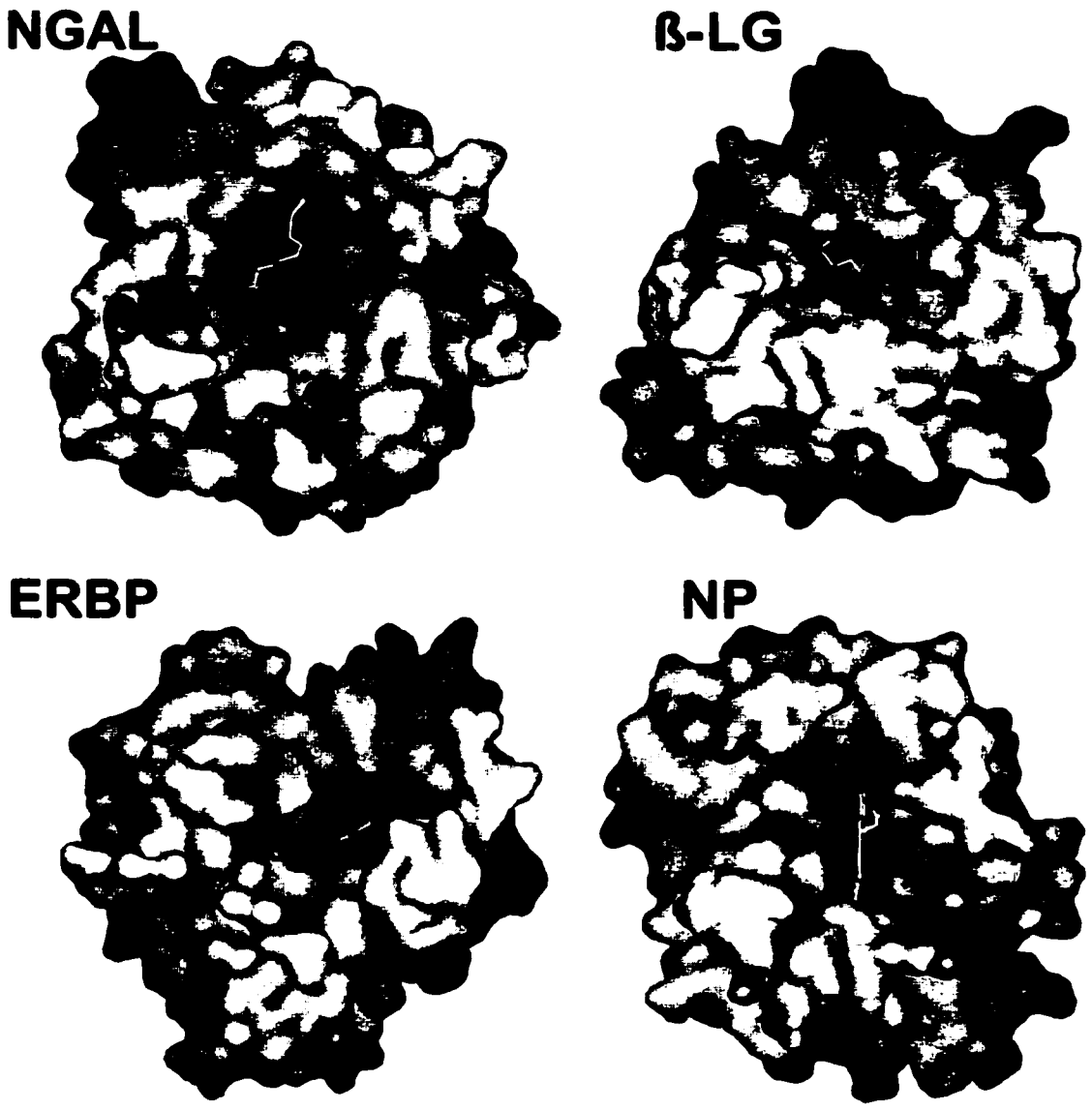


Figure 10: Lipocalin Binding Sites. Molecular surfaces colored by electrostatic potential (blue, positive; red, negative) for NGAL, ß-LG, ERBP, and NP. The bound ligands are shown as stick models: NCA in NGAL, palmitate in ß-LG, retinoic acid in ERBP, and heme in NP. The ligand atoms are colored as in Figure 4.

Chapter 3: The Ligand for Neutrophil Gelatinase-associated Lipocalin

Bacterial Expression Results and Initial Chromphore Characterization

To produce NGAL protein more easily and cheaply than was possible with the baculovirus expression system for subsequent biochemical and structural studies, we expressed NGAL heterologously in XL1-Blue (stratagene) *E. coli* as a GST fusion protein as previously described[49]. Briefly, the NGAL-GST fusion expression plasmid was transfected into *E. coli* XL1Blue and plated on selective media. Ampicillin resistant colonies were picked and inoculated into liquid cultures, and incubated at 37°C with shaking until mid-log phase. Cultures were induced at mid-log phase with 1mM isopropylthiogalactose. The fusion protein was purified on glutathione-sepharose (Pharmacia) using the manufacturer's recommended protocol. We found that NGAL co-purified with a red chromophore characterized by U.V./ Vis absorption maxima at 334 and 498 nm (Figure 11). The protein was > 95% pure as judged by acrylamide gel electrophoresis (data not shown). The NGAL:chromophore complex was found to be stable over a wide range of pH, and in a mixture of 2% SDS and 8M urea for over 24 hours as assayed by gel-filtration chromatography. Positive-ion-mode electro-spray ionization and matrix-assisted laser desorption mass spectroscopy, however, suggested that the complex was not covalent (data not shown), but failed to provide an estimate for the mass of the ligand. The complex was separable by boiling in 8M urea followed by native acrylamide gel electrophoresis(data not shown). The migration of the chromophore at the dye front in 20% acrylamide indicated a large negative charge to mass ratio for the ligand. These data suggested that NGAL, a human protein, had acquired a negatively charged bacterial ligand.

The NGAL:chromophore complex was crystallized as described previously for monomeric NGAL[56] and 3.4 Å Cu-Kα diffraction data were collected. Previously described monomer crystals contained two ordered monomers in the asymmetric unit; however, in this case a crystal dehydration step in which the crystals were soaked in a

saturated ammonium sulfate glycerol solution prior to data collection resulted in three ordered molecules in the asymmetric unit (referred to as molecules A, B, and C) and a 1% reduction along the c-edge of the unit cell. The initial F_o-F_c electron density map revealed a large peak of positive electron density in the calyx centered between the guanidinium group of R87 and the Nε-atoms of K125 and K134. The location of the electron density peak confirmed that the chromophore was indeed specifically associated with the calyx of NGAL. The proximity of the electron density peak to positively charged lysine and arginine residues suggested that the ligand was negatively charged. Furthermore, a native anomalous difference electron density map revealed two significant peaks: one peak, with a height of 5.7σ, corresponded to the expected signal from the single disulfide bond in NGAL; the other peak, with a height of 7.2σ, was centered in the calyx at approximately the same position as the center of the F_o-F_c difference peak (Figure 12A). The height of the anomalous peak in the calyx indicated the presence of a heavy atom.

The identification of the heavy atom as iron was accomplished by graphite furnace atomic absorption spectroscopy (AA) and X-ray fluorescence spectroscopy. The molar ratio of iron to protein according to AA was variable among different protein preparations yet never exceeded 1:1 (Figure 12B). This ratio and the presence of a single large peak in the anomalous difference electron density map demonstrated that each molecule of NGAL binds a single iron atom in its calyx. In addition, spectroscopic and AA measurements demonstrated that >99% pure apo-NGAL can be prepared from *E. coli* BL21 (Figures 11A, 11B, and 12B).

The Role of Iron in Homeostasis, Infection and Neoplasia

Iron is an essential nutrient for nearly all life. [68] It has long been noted that supplementing iron to patients during bacterial infection increases the likelihood of a negative clinical outcome[69]. To limit the growth of microorganisms in host tissues, mammals utilize three proteins to minimize free iron: transferrin, lactoferrin, and ferritin.

Transferrin transports iron to and from host cells. Transferrin binding sites for iron are normally unsaturated leaving a capacity for binding of free iron[70]. Lactoferrin is a potent bacteriostatic agent first discovered in milk that is released by neutrophils at sites of inflammation[14]. Once released, lactoferrin binds free iron and directly inhibits the growth of microorganisms[14]. Ferritins are the primary intracellular iron storage proteins[71]. Together these three proteins, along with specific cellular receptors and regulatory proteins, keep the level of free iron in the body at approximately 10^{-24}M[72].

The binding of iron to NGAL suggested that NGAL might function analogously to lactoferrin as an acute antibacterial agent by limiting the availability of free iron. However, both the large size and positive charge of the NGAL calyx, and the size of the F_o-F_c electron density feature in the calyx are wholly inconsistent with the identification of the NGAL ligand as a positively charged iron(II) or iron(III) ion. Thus NGAL prepared from XL1-Blue *E. coli* was most likely bound to a negatively charged iron-containing molecule acquired from the bacteria in which it was expressed. Nitrophorin[73] is a lipocalin which binds heme as a co-factor in its calyx and is a close structural neighbor of NGAL. Therefore we considered whether NGAL might bind bacterial heme. Comparison of the calyxes of Nitrophorin and NGAL (Figure 10), however, shows marked differences with respect to shape and charge. Nitrophorin has a coin-slot shaped calyx with a slightly negative surface potential suitable for binding planar heme molecules. The tri-lobed NGAL calyx is both more bowl-shaped and positively charged, suitable for binding non-planar, negatively charged molecules. These differences in the calyxes suggested that the NGAL chromophore was not bacterial heme.

Siderophores

Bacteria have developed several mechanisms for obtaining iron from the large reservoir of bound iron in the mammalian host. One of the most well-understood mechanisms is the synthesis, secretion, and re-uptake of siderophores. Siderophores are small molecules secreted by prokaryotic, eukaryotic, and archeal microorganisms that

bind iron extremely tightly for the purpose of scavenging iron from the extracellular environment. Siderophores bind iron orders of magnitude more tightly than lactoferrin or transferrin and can out-compete the mammalian system for iron[74]. *E. coli* synthesize only two siderophores: the catecholate-siderophore enterobactin (Ent) [75] and the citrate-siderophore aerobactin[76], however, they are also capable of utilizing siderophores produced by other microorganisms. Ent exhibits the tightest affinity for iron of any known molecule with a K_d of $10^{-49}M$[77]. To facilitate uptake of these siderophores, bacteria possess efficient, high-affinity outer membrane receptors such as the *E. coli* receptor FepA[78].

In general siderophores can be divided into two broad classes based on the functional group used to chelate the iron. Catecholate siderophores like Ent utilize a 2,3-dihydroxybenzene motif, while hydroxamate-siderophores use a hydroxamic acid motif (. .). Because hydroxamate has a lower pKa than catecholate, hydroxamate based siderophores are less susceptible to pH induced disassociation of the iron-siderophore complex. Several pathogenic bacterium that are closely related to *E. coli* such as *V. cholera* (synthesizes vibriobactin), and *S. typhimurium* (synthesizes enterobactin) synthesize catecholate based siderophores that are at least qualitatively similar to enterobactin from *E. coli*. Namely, they all utilize three catecholates positioned with near perfect octahedral geometry to chelate iron. The differences between enterobactin and vibriobactin for example lie at the linking groups used to join the catecholate moieties together. For a general overview of siderophore structure and function please consult the following reference[74].

Identification and Characterization of NGAL Ligand

In order to examine whether NGAL might bind a siderophore, apo-NGAL was assayed for binding to ferric-aerobactin and ferric-Ent (FeEnt). Aerobactin was mixed with apo-NGAL, dialyzed extensively, and assayed for binding by AA. No detectable

binding was observed. FeEnt was mixed with apo-NGAL and assayed for binding by ultrafiltration through a 10kD cutoff membrane. The flow-through was monitored visually and spectroscopically. The results indicated a tight binding with an undetectable free ligand concentration by U.V. absorbance. The NGAL:FeEnt complex was identical to the NGAL:chromophore complex both visually and spectroscopically (Figures 11A and 11B). On the basis of these data, the NGAL ligand was identified as enterobactin (Figure 13).

The ability of NGAL to bind a bacterial siderophore suggested that NGAL is either released by neutrophils at sites of infection and inflammation as a strategy to limit the growth of microorganisms by interfering with the bacterial iron scavenging pathway, or that it functions as a sensing mechanism to detect infection. However, it is unclear how such a sensing mechanism would serve the host because NGAL is only secreted after the neutrophil has already been signaled that there is a bacterial infection. This proposes a much more direct anti-bacterial mechanism for NGAL than has been previously suggested. For NGAL to effectively act as a growth inhibitor of *E. coli* by binding and sequestering FeEnt from the enterobactin receptor FepA, it must have a comparable affinity. The K_d of FepA for FeEnt has been measured by fluorescence spectroscopy at 50 nM [79]and by filter-binding at 0.2 nM[80]. Tryptophan fluorescence quenching analysis of NGAL in the presence of enterobactin yielded a K_d of 0.41±0.11nM at 20° C (Figure 14A).Briefly, the intrinsic tryptophan fluorescence of NGAL was monitored as a function of the concentration subsaturation quantities of added ligand. The resulting datapoints were fitted to an equation for a one-binding site model by non-linear regression. This demonstrates that NGAL has a sufficient affinity to effectively neutralize the use of enterobactin by pathogenic *E. coli* as an iron acquisition strategy by competing with FepA for binding to FeEnt.

A crystal structure of FeEnt has never been reported, presumably because the rapid oxidative breakdown of the molecule into dihydroxybenzoylserine (DHBS) [81] (Figure 13B) precludes crystallization. NGAL appears to have a rather dramatic protective effect on the breakdown of FeEnt *in vitro*. FeEnt was freshly prepared, a portion was used to make the FeEnt complex with NGAL and a portion was reserved. Breakdown of FeEnt was assayed in both solutions spectroscopically after one month to look for degradation-induced red-shifting of the visible absorbance peak (data not shown). No breakdown of the NGAL:FeEnt complex was observed, while free FeEnt exhibited complete breakdown. Furthermore, ligand extracted from month-old NGAL:FeEnt complex was spectroscopically identical to freshly prepared FeEnt. These results indicate that while iron bound to Ent *in vivo* is released by degradation of labile bonds in the triserine-trilactone backbone[82], the iron in the NGAL:FeEnt complex is likely to be relatively unavailable.

The Structure of the NGAL:FerricEnterobactin Complex

In order to understand the structural basis of ligand specificity of NGAL for FeEnt, we crystallized the NGAL:chromophore complex, collected synchrotron data to improve resolution and data quality, and determined the structure. The resulting structure at 2.4 Å with three molecules in the asymmetric unit shows that NGAL interacts with the catecholate moieties immediately surrounding the iron, and that the triserine-trilactone backbone of enterobactin is not a major determinant of ligand specificity for NGAL (Figures 15 and 16).

The NGAL:FeEnt structure is essentially identical to the previously described NGAL crystal structures. The largest mainchain rmsd among all six protein molecules

from the three crystal structures is 0.76 Å. The largest Cα-Cα distance among all six structures is 3.3 Å for P101 between the structure of dimeric NGAL and molecule C of this structure. P101 participates in a crystal contact in the dimer crystal, and lies at the tip of an extended loop. The local difference at this residue is not propagated to other residues within the loop. The second (K73, 2.7 Å) and third (K46, 2.5 Å) largest differences also lie in loops and are distant from the calyx. The lack of conformational changes between the three structures suggests that NGAL is quite rigid and does not undergo notable conformational changes due to ligand binding, ionic strength of the solvent, or pH.

The total buried surface area of FeEnt in the calyx ranges from 461 $Å^2$ (62% of the total surface area of the molecule) to 520 $Å^2$ (70%) among the three protein molecules in the asymmetric unit. The total surface area of the protein buried by the ligand ranges from 364 $Å^2$ (3.8%) to 414 $Å^2$ (4.5%). The range in buried surface area is due to the state of residue W79. In molecule B, which exhibits the lowest buried surface area, the tryptophan is disordered and could be confidently modeled only as an alanine. In previous crystal structures, this tryptophan was also disordered except when interacting with a bound sulphate ion. Although NGAL buries approximately two-thirds of the FeEnt molecule, large sections of the NGAL calyx remain unfilled by the ligand (Figure 15B). This suggests that NGAL may also be capable of binding other larger siderophores, and may represent a general mechanism for inhibiting the growth of catecholate-siderophore-producing microorganisms.

FeEnt binds NGAL with each catecholate moiety in a distinct pocket of the NGAL calyx (Figure 15A). Residues contacting the FeEnt ligand are shown in Figure 16A. With the exception of W79 and R81, the residues that interact with FeEnt are in equivalent positions in all three molecules in the asymmetric unit (Figure 16B). W79 occupies the equivalent position at the boundary between Pockets 2 and 3 in molecules A and C, but is disordered and modeled as an alanine in molecule B. The side chain of R81 adopts two different conformations among the three molecules in the asymmetric unit. In molecules A and C, the side chain of R81 forms a hydrogen bond to Oγ of S68, and helps form the floor of Pocket 2. In molecule B, there is a 150° rotation around the χ1 torsion

angle. This movement orients the side chain such that it forms the boundary between Pockets 2 and 3 and takes the place of the disordered W79 side chain.

NGAL makes the most extensive contacts with FeEnt in Pocket 1. In Pocket 1, the catecholate lies sandwiched between the side chains of K134 and K125, under Y132 and the loop formed by G38-I41, rests on top of the side chains of Y106 and F123, and interacts with the mainchain atoms of K124 and F133, which form the back wall of the pocket. For the catecholate moiety in Pocket 1, the carbonyl oxygen of the amide forms a water-mediated hydrogen bond to the backbone nitrogens of N39 and A40 as well as the Oδ1-atom of N39. Similarly, the O2-atom in Pocket 1 forms a hydrogen bond to the hydroxyl of Y106. The other pockets contact the enterobactin molecule much less extensively (Figures 15A and 16A). Pocket 2 is defined by the Nε–atom of K134, and the side chains of Y52, L36, T54, S68, L70, W79 (molecules A and C), R81 (molecules A, B, and C), and the hydroxyl of Y106. Pocket 3 is defined by the side-chain of K125, the guanidinium group of Arg 81 (molecules A and B), W49 (molecules A and C), and the Y100-L103 loop.

Some of the protein-ligand interactions seen in this structure are unique and have not been previously reported. The Nε-atom of K125 lies within 4.3 Å of the center of the catecholate in Pocket 1 and 3.4 Å of the center of the catecholate in Pocket 3 (Figure 16C). Similarly, the Nε-atom of K134 lies within 4.6 Å of the catecholate ring in Pocket 1 of the NGAL calyx and 3.3 Å of the catecholate in Pocket 2. Finally, the Cζ-atom of R81 in molecule B lies within 5.1 Å and 5.3 Å of the catecholate groups in Pockets 2 and 3 respectively. These distances are within the range observed for favorable protein-cation-π interactions[83].

Cation-π interactions in proteins are an interaction between the positive charge of lysine or arginine side-chains and the quadrupole moment associated with the delocalized π-electrons of an aromatic functional group such as tryptophan, tyrosine or phenylalanine[84]. Semi-empirical calculations using the small molecule crystal structure of vanadium-enterobactin[85] as a starting model for FeEnt demonstrate the highly delocalized nature of the net negative charge (Figure 16D). Thus, the NGAL:FeEnt interactions with lysine and arginine side chains are a hybrid between a simple ionic

interaction and a cation-π interaction in that both interacting moieties have a net charge. These interactions are similar to those seen in the binding of phosphotyrosine to an SH2 domain[86], which is also a complex in which an arginine and a lysine interact with a negatively charged ligand. However, in that complex the actual aromatic ring remains uncharged.

The structure of the hydroxamate-siderophore ferrichrome in complex with FhuA[87] represents one of two other examples of an ordered siderophore-protein complex. The binding pocket of FhuA and the calyx of NGAL both contain a number of aromatic residues. The interactions mediated by the aromatic residues in the two siderophore binding proteins are, however, different because of the neutral charge of the FhuA ligand ferrichrome as compared with the negative charge of FeEnt. Nevertheless, the aromatic residues of both FhuA and NGAL are involved in van der Waals interactions with their ligands as well as tyrosine-hydroxyl-mediated hydrogen bonds[88]. FhuA, however, cannot form cation-π interactions with its ligand because ferrichrome lacks aromatic functional groups. This may explain the 244-fold lower affinity of FhuA for ferrichrome[89] compared to the affinity of NGAL for FeEnt.

The structure of the *E. coli* outermembrane receptor FecA in complex with ferric-citrate has recently been published by Ferguson *et al.* [90]. In their structure, there are a number of positively charged residues which form hydrogen bonds and favorable ionic interactions with the negatively charged ferric-citrate ligand. The binding pocket shares superficial similarities with the calyx of NGAL due to the positive charges, however, the spatial arrangement of the positively charged residues is quite different between the two structures. There is no structural or sequence homology between the NGAL calyx and the FecA binding site.

In the crystal structure of the FepA:FeEnt complex[79] both the binding pocket and the ligand are disordered. The authors conclude that a conformational change induced by ligand binding must take place that is inhibited in the crystal. This is supported by the fact that there are no regions in the FepA structure with apparent structural homology to the NGAL calyx. However, like NGAL, the crystal structure and biochemical data have implicated several positively charged and aromatic residues of FepA in the binding of

FeEnt[91]. Therefore, it is likely that NGAL and FepA share similarities in the way in which they recognize FeEnt.

After several rounds of model building and refinement, it was evident that the triserine-trilactone backbone of the ligand was at least partially degraded as exhibited by large negative F_o-F_c difference peaks around the backbone atoms after refinement in CNS (Figure 17A) and by unreasonably high libration tensors for the FeEnt molecules following TLS refinement in refmac5. In spite of the aforementioned protective effect NGAL has on enterobactin stability, the extent of ligand degradation was not surprising due to the seven months required to grow diffraction-quality crystals. The lack of conformational flexibility of enterobactin allowed us to position the intact ligand with a high degree of confidence despite the large negative F_o-F_c peaks overlapping ligand atoms. The NGAL:FeEnt model was refined to an R_{work} and R_{free} of 23.8 and 28.0 respectively.

Although we finished refinement with a model of degraded enterobactin, the resulting structure is nearly identical to the structure with the intact ligand. The rmsd calculated on all atoms of residues contacting FeEnt between the two structures is 0.3 Å. The R_{work} and R_{free} values for our model of NGAL:degradedFeEnt are 1.1% and 0.9% better respectively compared to the NGAL:FeEnt model. This difference is mostly due to the improved agreement of our model of the degraded ligand to the crystallographic data (Figure 17B). With one exception described below, the interactions between NGAL and either FeEnt or the model of the degraded ligand are the same. The unambiguous position of the iron atom in the electron density maps, the rigidity of FeEnt molecule, the small difference in the R_{work} and R_{free} values between the two structures, and the reasonable agreement between the positions of the catecholate moieties of the intact FeEnt molecule and the electron density maps demonstrates the validity of the NGAL:FeEnt model. Data collection and refinement statistics for the NGAL:degFeEnt and NGAL:FeEnt model are presented in Table 2.

Table 2: Data Collection and Refinement Statistics

Data Collection	
Resolution (Å)	20-2.4 (2.44-2.4)
Observations	312664

Table 2 Continued

Unique reflections	30746	
Completeness (%)	100 (100)	
I/σ(I)	37.3 (3.8)	
R_{sym} (%)	5.7 (51)	
Cell dimensions (Å) (P4(1)2(1)2)	115.1 115.1 115.2	
Refinement	NGAL:FeDHBx	NGAL:FeEnt
Resolution (Å)	19.76-2.4 (2.46-2.4)	19.76-2.4 (2.46-2.4)
Reflections (F>0)	28156	28156
Protein atoms (3 molecules/a.s.u)	4222	4185
Ligand atoms (including alternate conformations)	143	146
Solvent atoms	111 (2 SO_4^{-2},7 glycerol)	89
R_{work} (%)	22.5 (25.0)	23.8 (26.7)
R_{free} (%) (1470 reflections)	27.1 (35.9)	27.9 (36.7)
RMSD from ideality: bonds (Å)	0.013	0.013
Bond angles (°)	2.0	1.8
Average B-factor	37.1	33.8
Ramachandran plot: most favored (%)	87	85
Ramachandran plot: disallowed	0	2

Numbers in parentheses indicate values for the highest resolution shell

Degraded FeEnt was modeled as dihydroxybenzoate (DHBA) in Pockets 2 and 3, and dihydroxybenzoylserine in Pocket 1 complexed with iron; we refer to this mixture as FeDHBx (Figure 17C). In Pocket 1, all three molecules in the asymmetric unit superpose with 0.5 Å rmsd between matched atoms. In Pocket 2 of molecule A, two alternate conformations were modeled. In the first conformation, there is an apporiximately 42° right-handed rotation of the catecholate around the C5 to C2 vector relative to the intact enterobactin. The second conformation consists of a 160° flip of the catecholate in the left-handed direction around the same vector. This flip orients the carbonyl oxygen of the amide to within 1.3 Å of the previously identified carboxylate binding pocket of NGAL, and yields a hydrogen bond to the side chain of R81. This flip represents the one significant difference between the NGAL:FeEnt and NGAL:FeDHBx interactions. There is a 17° right-handed rotation of the catecholate in Pocket 2 of molecule B relative to the intact enterobactin along the vector from C2 to C6. In molecule C, the catecholate in Pocket 2 is in the flipped orientation as in molecule A. In Pocket 3 of molecule A, there is a 17° right-handed rotation of the catecholate along the vector from C1 to C4. In

molecule C two conformations were modeled for the DHBA in Pocket 3. The first conformation essentially superimposes the ligand in the NGAL:FeEnt model with a 0.2 Å rmsd for matched atoms. The second conformation consists of a 62° left-handed rotation about the C2 to C6 vector.

Because *E. coli* and other microorganisms also secrete the iron-chelating compound DHBA under iron limiting conditions, and our model of the degraded FeEnt ligand suggested that NGAL might bind ferric-DHBA (FeDHBA), we measured the strength of the NGAL:FeDHBA interaction by tryptophan fluorescence quenching (Figure 14B). The K_d of 7.9 ± 1.8 nM, while significantly higher than the K_d for FeEnt, is easily understood by taking into account the increased flexibility of the FeDHBA and the concomitant increased entropic cost associated with binding.

Implications of the NGAL structure on Function

Controversy concerning whether Ent chelates iron solely via the catechol oxygens (the catecholate mode of binding), or by the ortho-hydroxyl oxygen and carbonyl oxygen of the amide (salicylate mode of binding) [74,92,93]continues to rear its head in the literature. It is certainly true that at low pH in a non-aqueous environment Ent becomes protonated, which favors the salicylate mode of binding. However, our crystal structure at neutral pH unambiguously reveals the iron bound in the catecholate mode (Figure 17). While the ionic strength in the crystal is too high to be considered physiological, the structure suggests that either NGAL may alter the first protonation constant to stabilize the catecholate mode of binding over the salicylate mode, or that NGAL preferentially recognizes and binds to FeEnt complexes in the catecholate mode, or that ferric-enterobactin in fact only binds in the catecholate mode . This is significant because it is suggested that protonation of FeEnt to promote the salicylate mode of binding is part of a secondary pathway used by bacteria for removal of iron by reduction at the outermembrane[94]. NGAL may stabilize the unprotonated, harder-to-reduce, form of FeEnt over the protonated, easier-to-reduce form and thus effectively sequester the iron

in the FeEnt complex. However it is perhaps more likely that enterobactin simply does not every chelate iron in the salicylate mode under physiological conditions.

Although many microorganisms secrete DHBA in response to low iron conditions, it has been assumed that this compound is not a siderophore[74] because an organism's ability to synthesize DHBA is not associated with increased virulence *in vivo*[95]. However, our results suggest that DHBA is not associated with increased virulence *in vivo* due to NGAL, which binds to and sequesters FeDHBA. Similarly, the biosynthesis of Ent itself is not a virulence factor for *E. coli* in mammals. In contrast, the synthesis of the citrate-siderophore aerobactin, which has lower affinity for iron, is a virulence factor[96]. Our results provide a model to explain this puzzling phenomenon: *in vivo* NGAL acts to prevent Ent and DHBA-mediated iron acquisition by *E. coli*, while no such defense mechanism has been discovered that counteracts aerobactin-mediated iron acquisition.

To test the hypothesis that NGAL interferes with iron uptake by bacteria, we investigated the ability of NGAL to inhibit the growth of bacteria in defined media. *E. coli* XL1Blue transfected with the NGAL expression plasmid failed to grow in M9 minimal media unless supplemented with 10 μM iron. Furthermore, the addition of endogenous, purified, recombinant, 5 μM apo-NGAL to M9 media resulted in a 20-fold inhibition of the growth of *E. coli* strain XL1-Blue (Figure 18A) which was completely rescued by the addition of excess iron (Figure 18B). Although iron is not deliberately added to M9 media, due to the ubiquity of trace iron contamination in reagents and glassware the concentration of available iron in M9 has been independently reported to be 2 μM[97] and 0.1 μM[98]. This is sufficient iron for *E. coli* to grow to saturation if siderophore metabolism proceeds unimpeded. NGAL, however, interferes with uptake of ferric siderophore from the extracellular medium by sequestering siderophores away from the *E. coli* receptor. The rescue of the growth inhibition effect by excess iron demonstrates that NGAL does not directly interfere with a membrane associated iron transport mechanism such as a pore or a receptor, or have some additional antibacterial function.

NGAL plays a role in a variety of processes that appear unrelated to its antibacterial activity such as apoptosis[38] and cell differentiation[42]. The identification of the bacterial ligand, the crystal structure of the complex, and elucidation of the determinants of ligand specificity provide insight into NGAL's other functions. There have been several reports of unidentified, mammalian, low-molecular-weight, siderophore-like growth-factors[99] and iron-transport factors[100-102]. Furthermore, recent experiments show that recombinant NGAL, is capable of binding, transporting, and supplying iron to mouse cells *in vitro* (J. Barasch, in press). Together these data suggest that an unidentified, siderophore-like, mammalian ligand for NGAL may exist. The key requirement of such a ligand for recognition by NGAL is that it should utilize catecholate or structurally similar moieties to chelate iron.

There is confusion about whether NRL and 24p3 represent the functional rat and mouse homologues of NGAL[38]. An alignment of these three proteins with residues contributing to the NGAL-ligand interface identified as primary or secondary liganding residues is shown in Figure 19. Primary liganding residues contain at least one atom that has greater than 3 $Å^2$ buried by the ligand and is at most 5 Å distant from the ligand. Secondary liganding residues do not fit the criteria for primary liganding residues, yet contain at least one atom that is buried by the ligand. For primary residues, there have been two conservative substitutions: I41 in NGAL becomes valine in NRL and 24p3, while L103 becomes isoleucine and valine in NRL and 24p3 respectively. There has also been one non-conservative substitution at Q49, which becomes phenylalanine in both NRL and 24p3. For secondary residues, G38 becomes alanine in NRL and K124 becomes arginine and glutamate in 24p3 and NRL respectively. The overall sequence identity between NGAL and the mouse and rat homologues is 62% and 63.5% respectively.

The amino acid substitutions in the primary and secondary liganding residues of the rat and mouse homologues were modeled by mutating the respective amino acids in the NGAL structure followed by energy minimization in refmac5 with no X-ray restraint. The resulting structures were essentially identical to the starting model. The largest mainchain distance between the wild-type structure and the mouse and rat models is 0.19 Å between the carbonyl oxygen of residue 103 in the wild-type structure and both

models. There were no bad contacts between mutated residues and the ligand. The lack of differences in the calyx after modeling suggests that NGAL, NRL, and 24p3 have essentially identical calyxes. Therefore we presume that they bind the same or similar ligands as NGAL, and are functionally equivalent. These results are supported by functional studies of human NGAL in a mouse tissue culture cell-differentiation assay (J. Barasch, personal communication).

Materials and Methods

Data Collection and Structure Determination

Cu K-α X-ray data were collected from 20 Å to 3.4 Å with >99% completeness in all resolution bins and >10-fold redundancy. Anomalous difference electron density maps were calculated in CNS[61]. Crystallographic data were also collected at the advanced light source (ALS); Berkeley, CA beamline 5.0.1 to 2.4 Å, and processed using HKL2000[57] (Table 1).

The structure was solved by molecular replacement with the program EPMR[103] using the previously published structure of NGAL (PDB code: 1QQS) as the search model.

The NGAL:FeEnt structure was refined in CNS by simulated annealing with torsional dynamics followed by positional and individual B-factor refinement to an R_{work} and R_{free} of 0.28 and 0.33 respectively. Solvent molecules were added manually and the structure was then refined in Refmac5 using TLS parameter refinement[104,105] followed by positional and individual B-factor refinement to an R_{work} and R_{free} of 0.238 and 0.279 respectively. Six TLS groups were refined: one for each protein molecule and one for each ligand molecule in the asymmetric unit.

The NGAL:FeDHBx model was refined in Refmac5 using TLS parameter refinement followed by positional and individual B-factor refinement. Care was taken to maintain the same R_{free} set throughout. Manual model rebuilding was carried out using the program XtalView[60]. Data and refinement statistics of the NGAL:FeDHBx and

NGAL:FeEnt structures are found in Table1. The structure of NGAL:FeDHBx was deposited in the protein databank with accession id 1L6M.

Biochemical analysis of the NGAL:ligand complex

Atomic absorption measurements were performed on a Perkin-Elmer 405 graphite furnace atomic absorption spectrophotometer. The manufacturer's recommended protocol for measurement of iron in aqueous samples was followed. Briefly, a sample was dried at 99° C in the furnace, charred at 1200° C, and atomized at 2500° C. Standard curves were generated by serial dilution of a reference iron standard (Acros Organics). X-ray fluorescence measurements were carried out at the ALS on beamline 5.0.2.

Tryptophan fluorescence quenching was measured on a Hitachi F-2500 fluorometer with 5nm slit widths in Tris-buffered-saline 7.4 supplemented with 5mM Tris(2-carboxyethyl)phosphine hydrochloride (Pierce) and 32ug/ml ubiquitin (Sigma). Receptor concentrations were 22nM for measurement of the FeEnt affinity, and 120nM for FeDHBA. Enterobactin was provided by Dr. Kenneth Raymond at the University of California Berkeley. Fluorescence values were corrected for dilution upon addition of titrant, ligand-absorbance, ligand-fluorescence, and photo-induced quenching of receptor. Fluorescence data were analyzed by non-linear regression analysis of fluorescence response versus free ligand concentration using a one-site binding model as implemented in DYNAFIT[106]. As negative controls, BSA was assayed for binding to FeEnt and hemin was assayed for affinity to NGAL. Neither control exhibited saturation binding up to micromolar ligand concentrations (data not shown).

Computational analysis of NGAL and ligand structures

Surface accessibility calculations were performed using the program NACCESS[107] with all parameters set to their default values. Semi-empirical calculations on FeEnt were carried out using the program SPARTAN[108] using the vanadium enterobactin structure as a starting model.

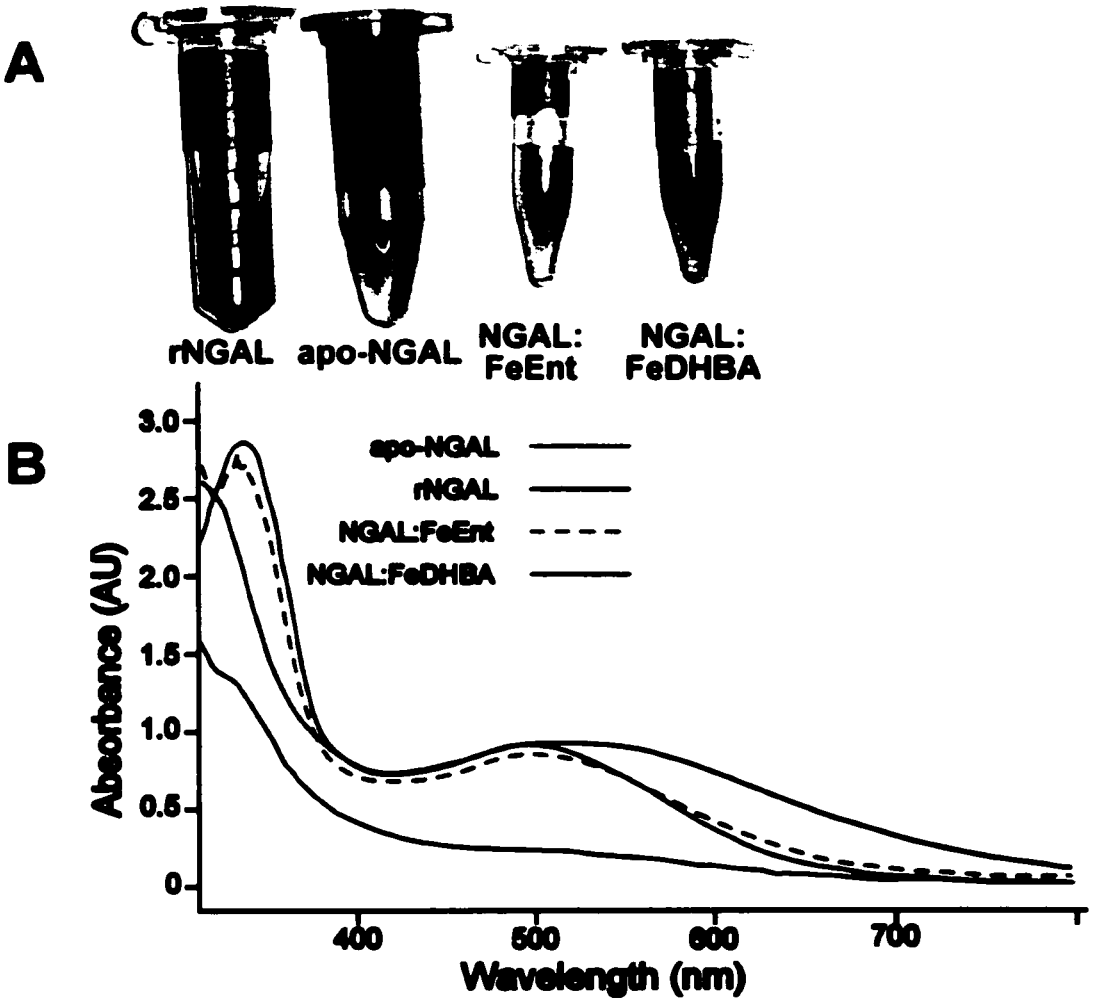


Figure 11: NGAL Binds Tightly to a Chromophore of Bacterial Origin (A) From left to right: NGAL from XL1-Blue at 10mg/ml, NGAL from *E. coli* BL21 at 91 mg/ml (apo-NGAL), and apo-NGAL saturated with ferric-enterobactin at 10mg/ml. NGAL with 2,3-dihydroxybenzoic acid at 5mg/ml is shown on the far right. (B) Spectroscopic data of NGAL ligand complexes; all samples are at 170μM.

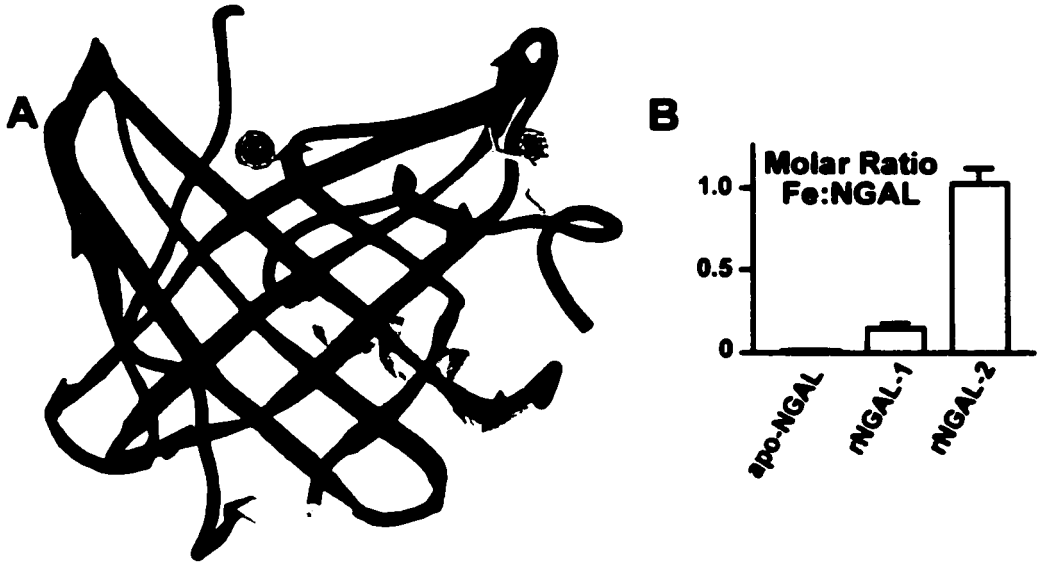


Figure 12: The NGAL Chromophore Contains Iron (A) 3.4 Å native anomalous difference electron density map of the NGAL:chromophore complex contoured in red and superimposed over a ribbon diagram of NGAL. (B) Molar ratio of iron to NGAL as determined by AA: apo-NGAL contains approximately 0.005 moles of iron per mole of NGAL; rNGAL-1, expressed in minimal media supplemented with 10μM iron, contains 0.17 moles of iron per mole of NGAL; and rNGAL-2 expressed in rich media, contains approximately 1 mole of iron per mole of NGAL.

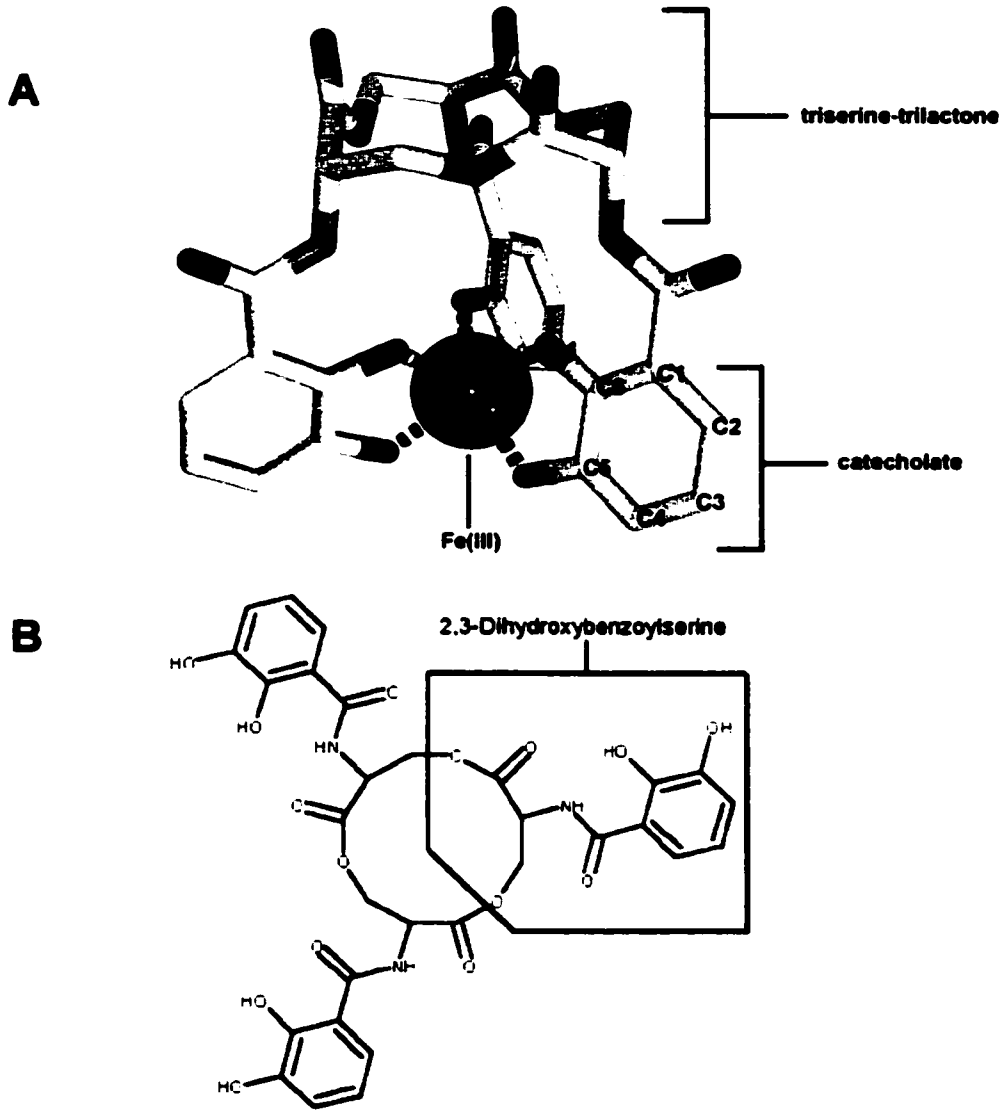


Figure 13: NGAL Binds the E. coli Siderophore Enterobactin (A) Structure of ferric-enterobactin color-coded by atom type (carbon, gray; nitrogen, blue; oxygen and iron, red). (B) Schematic diagram of deferri-enterobactin.

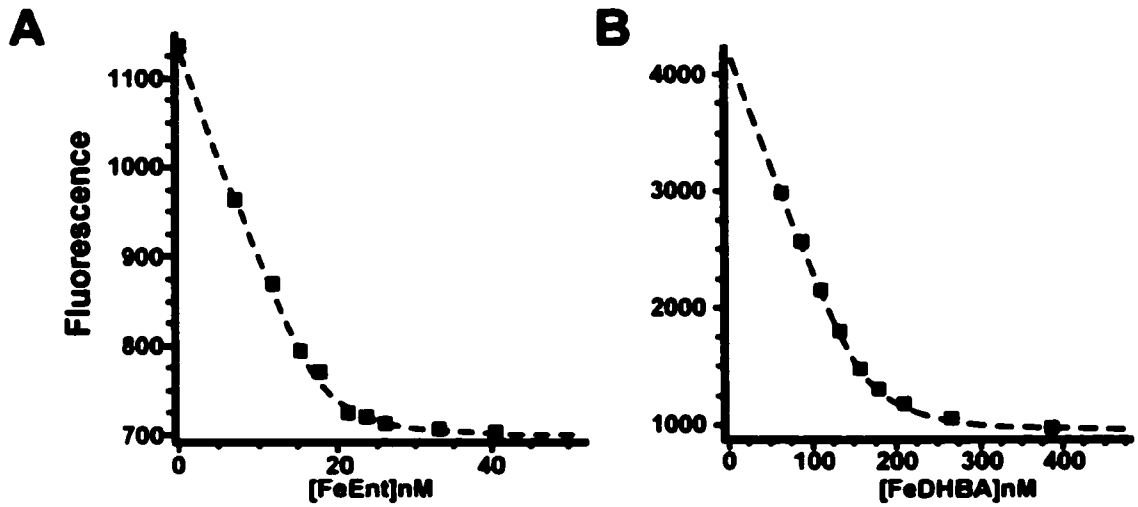


Figure 14: NGAL Binds Enterobactin with High Affinity. (A) Intrinsic tryptophan fluorescence quenching analysis of the NGAL:FeEnt interaction plotted as total fluorescence versus ligand concentration. Datapoints are indicated by filled squares; the dashed line indicates the best-fit line for a one-site binding model calculated by non-linear regression. (B) Tryptophan fluorescence quenching analysis of the NGAL:FeDHBA interaction.

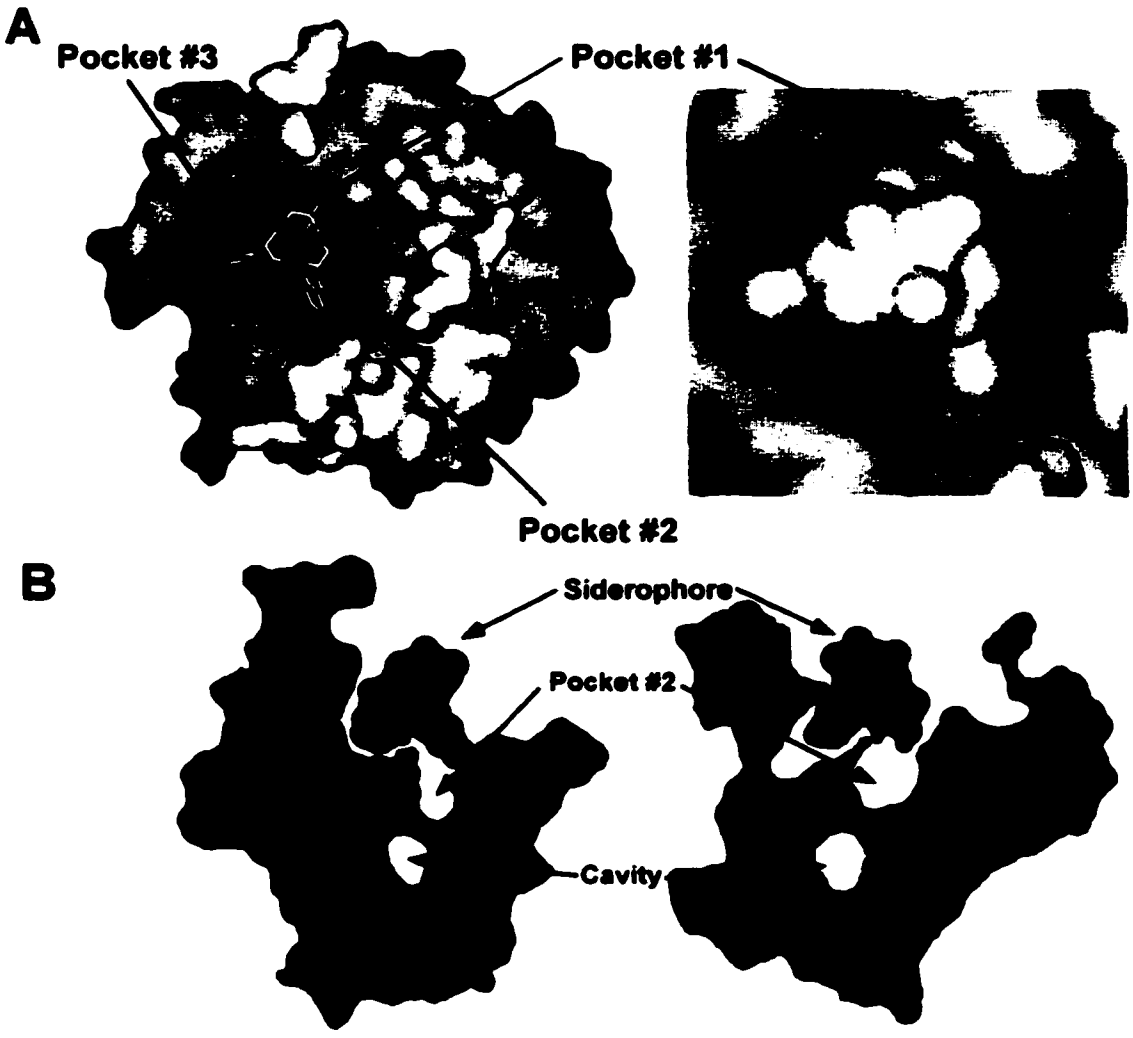


Figure 15: The NGAL:FeEnt interaction(A) Electrostatic surface potential of NGAL colored blue for positive and red for negative potential. The panel on the left shows a ball and stick model of enterobactin in the calyx colored by atom type. The panel on the right shows a close up view of the ligand in the calyx with the molecular surface of the ligand contoured in white. The individual pockets of the calyx are labeled as they are referred to in the text. (B) Orthogonal views of a thin slice of the calyx with the ligand cut roughly in half.

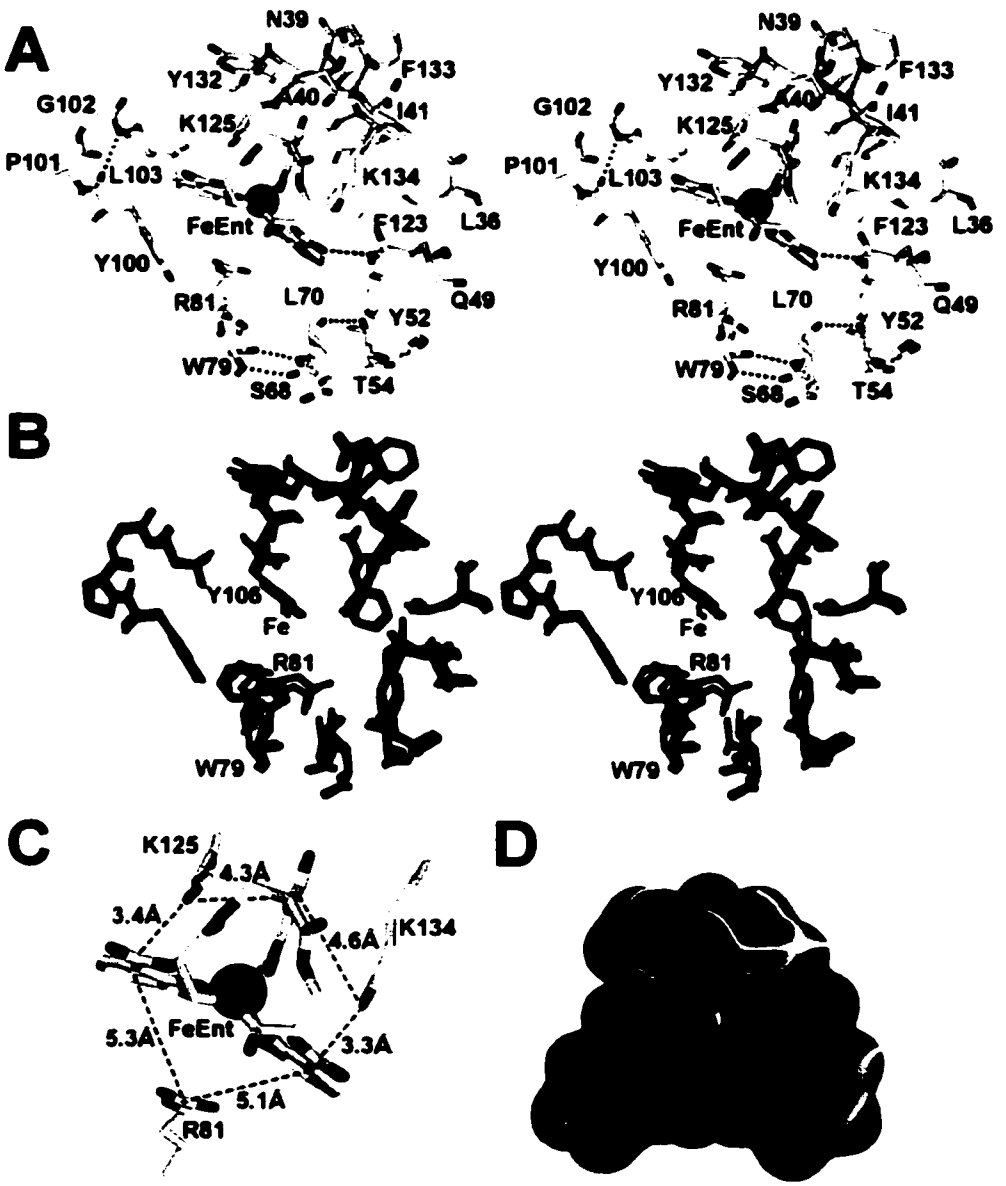


Figure 16: Closeup of the NGAL:FeEnt interaction. (A) Stereo view of the NGAL calyx showing residues that interact with FeEnt. (B) Stereo view of the NGAL calyx with all three molecules superimposed. Molecule A is colored blue, B purple, and C red. (C) Close up view of cationic protein residues of NGAL interacting with enterobactin. (D) Electrostatic potential of van der Waals surface of FeEnt calculated by semi-empirical methods. Red indicates most negative and blue is least negative potential.

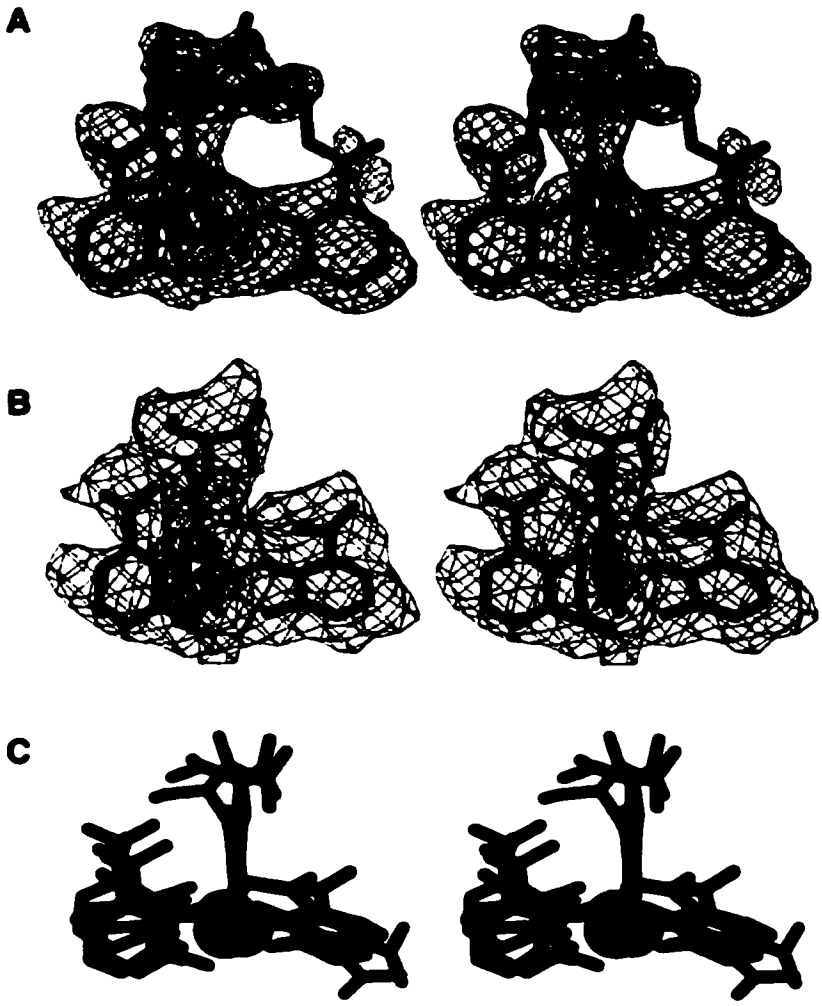


Figure 17: The Ligand in the Calyx of NGAL is Partially Degraded (A) 3σ F_o-F_c negative electron density contoured in red with 1σ $2F_o$- F_c electron density contoured in blue around FeEnt after refinement in CNS. (B) 1σ $2F_o$- F_c omit-refined electron density contoured in blue around degFeEnt with 3σ $2F_o$- F_c omit density contoured in green after refinement of TLS parameters in refmac5. (C) Stereo view of the three FeDHBx molecules in the asymmetric unit colored as described for Figure 16B.

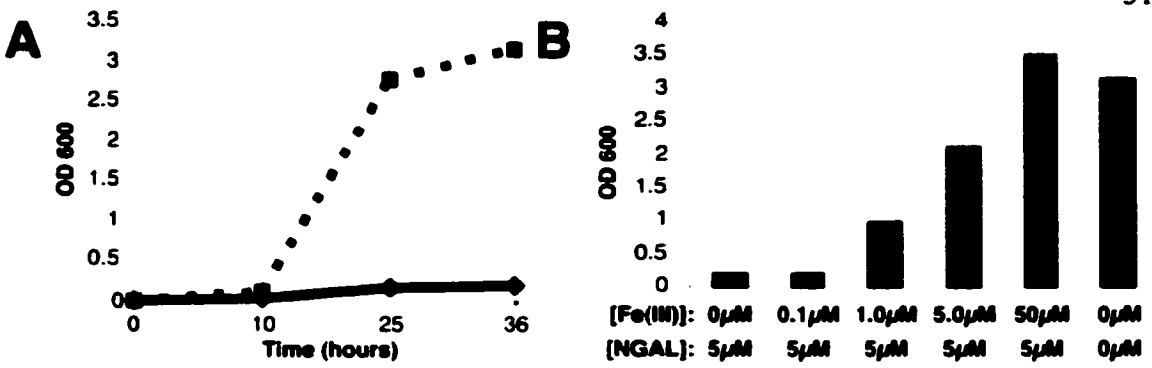


Figure 18. NGAL Limits the Growth of *E. coli* XL1-Blue in Iron-limited Minimal Media (a)The solid lines indicate growth in M9 medium with 5μM NGAL at 37° C. Dashed lines indicate growth in M9 media at 37° C. (b) E. coli XL1-Blue were grown in M9 media supplemented with 5μM NGAL and increasing concentrations of iron. Growth was measured by O.D. 600 at 36 hours at 37° C.

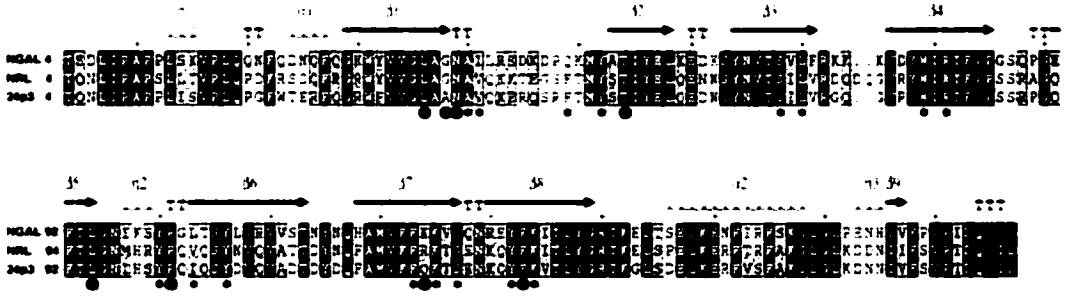


Figure 19. Sequence Alignment of 24p3 and NRL against NGAL generated in ESPript[109]: residues with stars denote first shell residues as defined in the text, dots denote second shell residues, residues shaded red are absolutely conserved, yellow denotes conservative substitutions. Secondary structural elements are shown above the sequence. Arrows indicate ß-strands, coils indicate α and 3_{10}-helices, and turns are indicated by a T. Every tenth residue is indicated by a dot.

End Notes

1. Mary, J.Y. Normal human granulopoiesis revisited. I. Blood data. *Biomed Pharmacother* **38**, 33-43 (1984).
2. Haddy, T.B., Rana, S.R. & Castro, O. Benign ethnic neutropenia: what is a normal absolute neutrophil count? *J Lab Clin Med* **133**, 15-22. (1999).
3. Mollinedo, F., Borregaard, N. & Boxer, L.A. Novel trends in neutrophil structure, function and development. *Immunol Today* **20**, 535-7. (1999).
4. da Silva, F.M., Massart-Leen, A.M. & Burvenich, C. Development and maturation of neutrophils. *Vet Q* **16**, 220-5. (1994).
5. Walker, R.I. & Willemze, R. Neutrophil kinetics and the regulation of granulopoiesis. *Rev Infect Dis* **2**, 282-92. (1980).
6. Keller, H.U. Motility, cell shape, and locomotion of neutrophil granulocytes. *Cell Motil* **3**, 47-60 (1983).
7. Keller, H.U. Shape, motility and locomotor responses of neutrophil granulocytes. *Agents Actions Suppl* **12**, 54-72 (1983).
8. Picker, L.J. Control of lymphocyte homing. *Curr Opin Immunol* **6**, 394-406. (1994).
9. Whited, S.C. & Gallin, J.I. Neutrophil chemotaxis. *Int J Dermatol* **19**, 130-8. (1980).
10. Harvath, L. Neutrophil chemotactic factors. *Exs* **59**, 35-52 (1991).
11. Sullivan, S.J., Daukas, G. & Zigmond, S.H. Asymmetric distribution of the chemotactic peptide receptor on polymorphonuclear leukocytes. *J Cell Biol* **99**, 1461-7. (1984).
12. Borregaard, N. Development of neutrophil granule diversity. *Ann N Y Acad Sci* **832**, 62-8 (1997).
13. Borregaard, N. *et al.* Human neutrophil granules and secretory vesicles. *Eur J Haematol* **51**, 187-98. (1993).
14. Ellison, R.T. The effects of lactoferrin on gram-negative bacteria. *Adv Exp Med Biol* **357**, 71-90 (1994).

15. Borregaard, N., Sehested, M., Nielsen, B.S., Sengelov, H. & Kjeldsen, L. Biosynthesis of granule proteins in normal human bone marrow cells. Gelatinase is a marker of terminal neutrophil differentiation. *Blood* **85**, 812-7. (1995).

16. Thomas, E.L., Lehrer, R.I. & Rest, R.F. Human neutrophil antimicrobial activity. *Rev Infect Dis* **10 Suppl 2**, S450-6. (1988).

17. Clark, R.A. Activation of the neutrophil respiratory burst oxidase. *J Infect Dis* **179 Suppl 2**, S309-17. (1999).

18. Fantone, J.C. & Ward, P.A. Polymorphonuclear leukocyte-mediated cell and tissue injury: oxygen metabolites and their relations to human disease. *Hum Pathol* **16**, 973-8. (1985).

19. Klebanoff, S.J. Oxygen metabolism and the toxic properties of phagocytes. *Ann Intern Med* **93**, 480-9. (1980).

20. Clifford, D.P. & Repine, J.E. Hydrogen peroxide mediated killing of bacteria. *Mol Cell Biochem* **49**, 143-9. (1982).

21. Ganz, T. *et al.* Defensins. Natural peptide antibiotics of human neutrophils. *J Clin Invest* **76**, 1427-35. (1985).

22. Baggiolini, M. The enzymes of the granules of polymorphonuclear leukocytes and their functions. *Enzyme* **13**, 132-60 (1972).

23. Triebel, S., Blaser, J., Reinke, H. & Tschesche, H. A 25 kDa alpha 2-microglobulin-related protein is a component of the 125 kDa form of human gelatinase. *FEBS Lett* **314**, 386-8. (1992).

24. Kjeldsen, L., Johnsen, A.H., Sengelov, H. & Borregaard, N. Isolation and primary structure of NGAL, a novel protein associated with human neutrophil gelatinase. *J Biol Chem* **268**, 10425-32 (1993).

25. Wang-Iverson, P., Pryzwansky, K.B., Spitznagel, J.K. & Cooney, M.H. Bactericidal capacity of phorbol myristate acetate-treated human polymorphonuclear leukocytes. *Infect Immun* **22**, 945-55. (1978).

26. Xu, S.Y. *et al.* Purification and characterization of a human neutrophil lipocalin (HNL) from the secondary granules of human neutrophils. *Scand J Clin Lab Invest* **54**, 365-76. (1994).

27. Kolkenbrock, H., Hecker-Kia, A., Orgel, D., Kinawi, A. & Ulbrich, N. Progelatinase B forms from human neutrophils. complex formation of monomer/lipocalin with TIMP-1. *Biol Chem* **377**, 529-33. (1996).

28. Flower, D.R., North, A.C. & Attwood, T.K. Mouse oncogene protein 24p3 is a member of the lipocalin protein family. *Biochem Biophys Res Commun* **180**, 69-74. (1991).

29. Chan, Y.L., Paz, V. & Wool, I.G. The primary structure of rat alpha 2 mu globulin-related protein. *Nucleic Acids Res* **16**, 11368. (1988).

30. Stoesz, S.P. & Gould, M.N. Overexpression of neu-related lipocalin (NRL) in neu-initiated but not ras or chemically initiated rat mammary carcinomas. *Oncogene* **11**, 2233-41. (1995).

31. Davies, B.R., Warren, J.R., Schmidt, G. & Rudland, P.S. Induction of a variety of preneoplasias and tumours in the mammary glands of transgenic rats. *Biochem Soc Symp* **63**, 167-84 (1998).

32. Hraba-Renevey, S., Turler, H., Kress, M., Salomon, C. & Weil, R. SV40-induced expression of mouse gene 24p3 involves a post-transcriptional mechanism. *Oncogene* **4**, 601-8. (1989).

33. Garay-Rojas, E., Harper, M., Hraba-Renevey, S. & Kress, M. An apparent autocrine mechanism amplifies the dexamethasone- and retinoic acid-induced expression of mouse lipocalin-encoding gene 24p3. *Gene* **170**, 173-80. (1996).

34. Davis, T.R., Tabatabai, L., Bruns, K., Hamilton, R.T. & Nilsen-Hamilton, M. Basic fibroblast growth factor induces 3T3 fibroblasts to synthesize and secrete a cyclophilin-like protein and beta 2-microglobulin. *Biochim Biophys Acta* **1095**, 145-52. (1991).

35. Kasik, J.W. & Rice, E.J. An increase in expression of the lipocalin 24p3 is found in mouse uterus coincident with birth. *Am J Obstet Gynecol* **173**, 613-7. (1995).

36. Liu, Q., Ryon, J. & Nilsen-Hamilton, M. Uterocalin: a mouse acute phase protein expressed in the uterus around birth. *Mol Reprod Dev* **46**, 507-14. (1997).

37. Meheus, L.A. *et al.* Identification by microsequencing of lipopolysaccharide-induced proteins secreted by mouse macrophages. *J Immunol* **151**, 1535-47. (1993).

38. Devireddy, L.R., Teodoro, J.G., Richard, F.A. & Green, M.R. Induction of apoptosis by a secreted lipocalin that is transcriptionally regulated by il-3 deprivation. *Science* **293**, 829-34. (2001).

39. Nielsen, B.S. *et al.* Induction of NGAL synthesis in epithelial cells of human colorectal neoplasia and inflammatory bowel diseases. *Gut* **38**, 414-20 (1996).

40. Cowland, J.B. & Borregaard, N. Molecular characterization and pattern of tissue expression of the gene for neutrophil gelatinase-associated lipocalin from humans. *Genomics* **45**, 17-23. (1997).

41. Bartsch, S. & Tschesche, H. Cloning and expression of human neutrophil lipocalin cDNA derived from bone marrow and ovarian cancer cells. *FEBS Lett* **357**, 255-9 (1995).

42. Zerega, B., Cermelli, S., Michelis, B., Cancedda, R. & Cancedda, F.D. Expression of NRL/NGAL (neu-related lipocalin/neutrophil gelatinase-associated lipocalin) during mammalian embryonic development and in inflammation. *Eur J Cell Biol* **79**, 165-72. (2000).

43. Friedl, A., Stoesz, S.P., Buckley, P. & Gould, M.N. Neutrophil gelatinase-associated lipocalin in normal and neoplastic human tissues. Cell type-specific pattern of expression. *Histochem J* **31**, 433-41 (1999).

44. Flower, D.R. Multiple molecular recognition properties of the lipocalin protein family. *J Mol Recognit* **8**, 185-95. (1995).

45. Flower, D.R., North, A.C. & Attwood, T.K. Structure and sequence relationships in the lipocalins and related proteins. *Protein Sci* **2**, 753-61. (1993).

46. Flower, D.R., North, A.C. & Sansom, C.E. The lipocalin protein family: structural and sequence overview. *Biochim Biophys Acta* **1482**, 9-24. (2000).

47. Flower, D.R. The lipocalin protein family: structure and function. *Biochem J* **318**, 1-14. (1996).

48. Bratt, T., Ohlson, S. & Borregaard, N. Interactions between neutrophil gelatinase-associated lipocalin and natural lipophilic ligands. *Biochim Biophys Acta* **1472**, 262-9. (1999).

49. Bundgaard, J.R., Sengelov, H., Borregaard, N. & Kjeldsen, L. Molecular cloning and expression of a cDNA encoding NGAL: a lipocalin expressed in human neutrophils. *Biochem Biophys Res Commun* **202**, 1468-75. (1994).

50. Allen, R.A., Erickson, R.W. & Jesaitis, A.J. Identification of a human neutrophil protein of Mr 24 000 that binds N- formyl peptides: co-sedimentation with specific granules. *Biochim Biophys Acta* **991**, 123-33. (1989).

51. Sengelov, H., Boulay, F., Kjeldsen, L. & Borregaard, N. Subcellular localization and translocation of the receptor for N-formylmethionyl-leucyl-phenylalanine in human neutrophils. *Biochem J* **299**, 473-9. (1994).

52. Chu, S.T., Lin, H.J. & Chen, Y.H. Complex formation between a formyl peptide and 24p3 protein with a blocked N-terminus of pyroglutamate. *J Pept Res* **49**, 582-5. (1997).

53. Jr., C.W.C. & Sweet, R.M. (eds.). *Macromolecular Crystallography Part A*, (Academic Press, San Diego, 1997).

54. Jr., C.W.C. & Sweet, R.M. (eds.). *Macromolecular Crystallography Part B*, (Academic Press, San Diego, 1997).

55. Kleywegt, G.J. & Brunger, A.T. Checking your imagination: applications of the free R value. *Structure* **4**, 897-904. (1996).

56. Strong, R.K. *et al.* Expression, purification, crystallization and crystallographic characterization of dimeric and monomeric human neutrophil gelatinase associated lipocalin (NGAL). *Acta Cryst* **D54**, 93-5 (1998).

57. Otwinowski, Z. & Minor, W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. in *Methods in Enzymology*, Vol. 276 (eds. Jr., C.W.C. & Sweet, R.M.) 307-326 (Academic Press, NY, 1997).

58. De La Fortelle, E. & Bricogne, G. Maximum-Likelihood Heavy-Atom Parameter Refinement for Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction Methods. *Methods Enzymol* **276**, 472-494 (1997).

59. Abrahams, J.P. & Leslie, A.G.W. Methods Used in the structure determination of bovine mitochondrial F_1 ATPase. *Acta Cryst* **D52**, 30-42 (1996).

60. McRee, D.E. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* **125**, 156-65. (1999).

61. Brunger, A.T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Cryst* **D54**, 905-921 (1998).

62. Kissinger, C.R., Gehlhaar, D.K. & Fogel, D.B. Rapid automated molecular replacement by evolutionary search. *Acta Cryst* **D55**, 484-91. (1999).

63. Reed, R.J. Improved Fourier coefficients for maps using phases from partial structures with errors. *Acta Cryst* **A42**, 140-149 (1986).

64. Morris, A.L., MacArthur, M.W., Hutchinson, E.G. & Thornton, J.M. Stereochemical quality of protein structure coordinates. *Proteins* **12**, 345-364 (1992).

65. Holm, L. & Sander, C. Touring protein fold space with Dali/FSSP. *Nucleic Acids Res* **26**, 316-9. (1998).

66. Coles, M., Diercks, T., Muehlenweg, B., Bartsch, S., Zolzer, V., Tschesche, H., Kessler, H.The solution structure and dynamics of human neutrophil gelatinase-associated lipocalin. *J Mol Biol* **289**, 139-57. (1999).

67. Stewart, J.M. The cytoplasmic fatty-acid-binding proteins: thirty years and counting. *Cell Mol Life Sci* **57**, 1345-59. (2000).

68. Neilands, J.B. Siderophores: structure and function of microbial iron transport compounds. *J Biol Chem* **270**, 26723-6. (1995).

69. Weinberg, E.D. Iron withholding: a defense against infection and neoplasia. *Physiol Rev* **64**, 65-102. (1984).

70. Ponka, P., Beaumont, C. & Richardson, D.R. Function and regulation of transferrin and ferritin. *Semin Hematol* **35**, 35-54. (1998).

71. Harrison, P.M. & Arosio, P. The ferritins: molecular properties, iron storage function and cellular regulation. *Biochim Biophys Acta* **1275**, 161-203. (1996).

72. Otto, B.R., Verweij-van Vught, A.M. & MacLaren, D.M. Transferrins and heme-compounds as iron sources for pathogenic bacteria. *Crit Rev Microbiol* **18**, 217-33 (1992).

73. Andersen, J.F., Weichsel, A., Balfour, C.A., Champagne, D.E. & Montfort, W.R. The crystal structure of nitrophorin 4 at 1.5 A resolution: transport of nitric oxide by a lipocalin-based heme protein. *Structure* **6**, 1315-27. (1998).

74. Ratledge, C. & Dover, L.G. Iron metabolism in pathogenic bacteria. *Annu Rev Microbiol* **54**, 881-941 (2000).

75. Pollack, J.R. & Neilands, J.B. Enterobactin, an iron transport compound from Salmonella typhimurium. *Biochem Biophys Res Commun* **38**, 989-92. (1970).

76. Gibson, F. & Magrath, D.I. The isolation and characterization of a hydroxamic acid (aerobactin) formed by Aerobacter aerogenes 62-I. *Biochim Biophys Acta* **192**, 175-84. (1969).

77. Loomis, L.D. & Raymond, K.N. Solution Equilibria of Enterobactin and Metal-Enterobactin Complexes. *Inorganic Chemistry* **30**, 906-911 (1991).

78. Wookey, P. & Rosenberg, H. Involvement of inner and outer membrane components in the transport of iron and in colicin B action in Escherichia coli. *J Bacteriol* **133**, 661-6. (1978).

79. Buchanan, S.K. *et al.* Crystal structure of the outer membrane active transporter FepA from Escherichia coli. *Nat Struct Biol* **6**, 56-63. (1999).

80. Newton, S.M., Igo, J.D., Scott, D.C. & Klebba, P.E. Effect of loop deletions on the binding and transport of ferric enterobactin by FepA. *Mol Microbiol* **32**, 1153-65. (1999).

81. O'Brien, I.G. & Gibson, F. The structure of enterochelin and related 2,3-dihydroxy-N-benzoylserine conjugates from Escherichia coli. *Biochim Biophys Acta* **215**, 393-402. (1970).

82. Hollifield, W.C., Jr. & Neilands, J.B. Ferric enterobactin transport system in Escherichia coli K-12. Extraction, assay, and specificity of the outer membrane receptor. *Biochemistry* **17**, 1922-8. (1978).

83. Gallivan, J.P. & Dougherty, D.A. Cation-pi interactions in structural biology. *Proc Natl Acad Sci U S A* **96**, 9459-64. (1999).

84. Dougherty, D.A. Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **271**, 163-8. (1996).

85. Karpishin, T.B., Dewey, T.M. & Raymond, K.N. The Vanadium (IV) Enterobactin Complex: Structural, Spectroscopic, and Electrochemical Characterization. *J Am Chem Soc* **115**, 1842-1851 (1993).

86. Waksman, G. Crystal structure of the phosphotyrosine recognition domain SH2 of the Src oncogene product complexed with tyrosine-phosphorylated peptides. *Cell Mol Biol (Noisy-le-grand)* **40**, 611-8. (1994).

87. Ferguson, A.D., Hofmann, E., Coulton, J.W., Diederichs, K. & Welte, W. Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. *Science* **282**, 2215-20. (1998).

88. Locher, K.P. *et al.* Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. *Cell* **95**, 771-8. (1998).

89. Locher, K.P. & Rosenbusch, J.P. Oligomeric states and siderophore binding of the ligand-gated FhuA protein that forms channels across Escherichia coli outer membranes. *Eur J Biochem* **247**, 770-5. (1997).

90. Ferguson, A.D. *et al.* Structural basis of gating by the outer membrane transporter FecA. *Science* **295**, 1715-9. (2002).

91. Cao, Z., Qi, Z., Sprencel, C., Newton, S.M. & Klebba, P.E. Aromatic components of two ferric enterobactin binding sites in Escherichia coli FepA. *Mol Microbiol* **37**, 1306-17. (2000).

92. Cass, M.E., Garrett, T.M. & Raymond, K.N. The Salicylate Mode of Bonding in Protonated Ferric Enterobactin Analogues. *J Am Chem Soc* **111**, 1677-1682 (1989).

93. Cohen, S.M., Meyer, M. & Raymond, K.N. Enterobactin Protonation and Iron Release: Hexadentate Tris-Salycilate Ligands as Models for Triprotonated Enterobactin. *J Am Chem Soc* **120**, 6277-6286 (1998).

94. Lee, C.-W., Ecker, D.J. & Raymond, K.N. The pH-Dependant Reduction of Ferric Enterobactin Probed by Electrochemical Methods and Its Implications for Microbial Iron Transport. *J Am Chem Soc* **107**, 6920-6923 (1985).

95. Bellaire, B.H., Elzer, P.H., Baldwin, C.L. & Roop, R.M., 2nd. The siderophore 2,3-dihydroxybenzoic acid is not required for virulence of Brucella abortus in BALB/c mice. *Infect Immun* **67**, 2615-8. (1999).

96. Wooldridge, K.G. & Williams, P.H. Iron uptake mechanisms of pathogenic bacteria. *FEMS Microbiol Rev* **12**, 325-48. (1993).

97. Roberts, R.B., Abelson, P.H., Cowie, D.B., Bolton, E.T. & Britten, R.J. Studies of biosynthesis in *Escherichia coli*. in *Publication 607* (Carnegie Institue of Washington, Washington D.C., 1963).

98. Klebba, P.E., McIntosh, M.A. & Neilands, J.B. Kinetics of biosynthesis of iron-regulated membrane proteins in Escherichia coli. *J Bacteriol* **149**, 880-8. (1982).

99. Fernandez-Pol, J.A. Isolation and characterization of a siderophore-like growth factor from mutants of SV40-transformed cells adapted to picolinic acid. *Cell* **14**, 489-99. (1978).

100. Jones, R.L., Peterson, C.M., Grady, R.W. & Cerami, A. Low molecular weight iron-binding factor from mammalian tissue that potentiates bacterial growth. *J Exp Med* **151**, 418-28. (1980).

101. Hershko, C., Graham, G., Bates, G.W. & Rachmilewitz, E.A. Non-specific serum iron in thalassaemia: an abnormal serum iron fraction of potential toxicity. *Br J Haematol* **40**, 255-63. (1978).

102. Craven, C.M. *et al.* Tissue distribution and clearance kinetics of non-transferrin-bound iron in the hypotransferrinemic mouse: a rodent model for hemochromatosis. *Proc Natl Acad Sci U S A* **84**, 3457-61. (1987).

103. Kissenger, C.R., Gehlhaar, D.K. & Fogel, D.B. Rapid automated molecular replacement by evolutionary search. *Acta Cryst* **D55**, 484-491 (1999).

104. Schomaker, V. & Trueblood, K.N. *Acta Cryst* **B54**, 507-514 (1998).

105. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Cryst* **D57**, 122-33. (2001).

106. Kuzmic, P. Program DYNAFIT for the Analysis of Enzyme Kinetic Data: Application to HIV Proteinase. *Anal. Biochem.* **237**, 260-273 (1996).

107. Hubbard, S.J. & Thornton, J.M. NACCESS. (Department of Biochemistry and Molecular Biology, University College, London, 1993).

108. Wavefunction. SPARTAN. 2.1 edn (, Ivine, CA, 1992).

109. Gouet, P., Courcelle, E., Stuart, D.I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305-8. (1999).

Bibliography

Abrahams, J. P., and Leslie, A. G. W. (1996). Methods Used in the structure determination of bovine mitochondrial F_1 ATPase. Acta Cryst *D52*, 30-42.

Allen, R. A., Erickson, R. W., and Jesaitis, A. J. (1989). Identification of a human neutrophil protein of Mr 24 000 that binds N- formyl peptides: co-sedimentation with specific granules. Biochim Biophys Acta *991*, 123-33.

Andersen, J. F., Weichsel, A., Balfour, C. A., Champagne, D. E., and Montfort, W. R. (1998). The crystal structure of nitrophorin 4 at 1.5 A resolution: transport of nitric oxide by a lipocalin-based heme protein. Structure *6*, 1315-27.

Baggiolini, M. (1972). The enzymes of the granules of polymorphonuclear leukocytes and their functions. Enzyme *13*, 132-60.

Bartsch, S., and Tschesche, H. (1995). Cloning and expression of human neutrophil lipocalin cDNA derived from bone marrow and ovarian cancer cells. FEBS Lett *357*, 255-9.

Bellaire, B. H., Elzer, P. H., Baldwin, C. L., and Roop, R. M., 2nd (1999). The siderophore 2,3-dihydroxybenzoic acid is not required for virulence of Brucella abortus in BALB/c mice. Infect Immun *67*, 2615-8.

Borregaard, N. (1997). Development of neutrophil granule diversity. Ann N Y Acad Sci *832*, 62-8.

Borregaard, N., Lollike, K., Kjeldsen, L., Sengelov, H., Bastholm, L., Nielsen, M. H., and Bainton, D. F. (1993). Human neutrophil granules and secretory vesicles. Eur J Haematol *51*, 187-98.

Borregaard, N., Sehested, M., Nielsen, B. S., Sengelov, H., and Kjeldsen, L. (1995). Biosynthesis of granule proteins in normal human bone marrow cells. Gelatinase is a marker of terminal neutrophil differentiation. Blood *85*, 812-7.

Bratt, T., Ohlson, S., and Borregaard, N. (1999). Interactions between neutrophil gelatinase-associated lipocalin and natural lipophilic ligands. Biochim Biophys Acta *1472*, 262-9.

Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr *D54*, 905-921.

Buchanan, S. K., Smith, B. S., Venkatramani, L., Xia, D., Esser, L., Palnitkar, M., Chakraborty, R., van der Helm, D., and Deisenhofer, J. (1999). Crystal structure of the outer membrane active transporter FepA from Escherichia coli. Nat Struct Biol *6*, 56-63.

Bundgaard, J. R., Sengelov, H., Borregaard, N., and Kjeldsen, L. (1994). Molecular cloning and expression of a cDNA encoding NGAL: a lipocalin expressed in human neutrophils. Biochem Biophys Res Commun *202*, 1468-75.

Cao, Z., Qi, Z., Sprencel, C., Newton, S. M., and Klebba, P. E. (2000). Aromatic components of two ferric enterobactin binding sites in Escherichia coli FepA. Mol Microbiol *37*, 1306-17.

Cass, M. E., Garrett, T. M., and Raymond, K. N. (1989). The Salicylate Mode of Bonding in Protonated Ferric Enterobactin Analogues. J Am Chem Soc *111*, 1677-1682.

Chan, Y. L., Paz, V., and Wool, I. G. (1988). The primary structure of rat alpha 2 mu globulin-related protein. Nucleic Acids Res *16*, 11368.

Chu, S. T., Lin, H. J., and Chen, Y. H. (1997). Complex formation between a formyl peptide and 24p3 protein with a blocked N-terminus of pyroglutamate. J Pept Res *49*, 582-5.

Clark, R. A. (1999). Activation of the neutrophil respiratory burst oxidase. J Infect Dis *179 Suppl 2*, S309-17.

Clifford, D. P., and Repine, J. E. (1982). Hydrogen peroxide mediated killing of bacteria. Mol Cell Biochem *49*, 143-9.

Cohen, S. M., Meyer, M., and Raymond, K. N. (1998). Enterobactin Protonation and Iron Release: Hexadentate Tris-Salycilate Ligands as Models for Triprotonated Enterobactin. J Am Chem Soc *120*, 6277-6286.

Coles, M., Diercks, T., Muehlenweg, B., Bartsch, S., Zolzer, V., Tschesche, H., and Kessler, H. (1999). The solution structure and dynamics of human neutrophil gelatinase-associated lipocalin. J Mol Biol *289*, 139-57.

Cowland, J. B., and Borregaard, N. (1997). Molecular characterization and pattern of tissue expression of the gene for neutrophil gelatinase-associated lipocalin from humans. Genomics *45*, 17-23.

Craven, C. M., Alexander, J., Eldridge, M., Kushner, J. P., Bernstein, S., and Kaplan, J. (1987). Tissue distribution and clearance kinetics of non-transferrin-bound iron in the hypotransferrinemic mouse: a rodent model for hemochromatosis. Proc Natl Acad Sci U S A *84*, 3457-61.

da Silva, F. M., Massart-Leen, A. M., and Burvenich, C. (1994). Development and maturation of neutrophils. Vet Q *16*, 220-5.

Davies, B. R., Warren, J. R., Schmidt, G., and Rudland, P. S. (1998). Induction of a variety of preneoplasias and tumours in the mammary glands of transgenic rats. Biochem Soc Symp *63*, 167-84.

Davis, T. R., Tabatabai, L., Bruns, K., Hamilton, R. T., and Nilsen-Hamilton, M. (1991). Basic fibroblast growth factor induces 3T3 fibroblasts to synthesize and secrete a cyclophilin-like protein and beta 2-microglobulin. Biochim Biophys Acta *1095*, 145-52.

De La Fortelle, E., and Bricogne, G. (1997). Maximum-Likelihood Heavy-Atom Parameter Refinement for Multiple Isomorphous Replacement and Multiwavelength Anomalous Diffraction Methods. Methods Enzymol *276*, 472-494.

Devireddy, L. R., Teodoro, J. G., Richard, F. A., and Green, M. R. (2001). Induction of apoptosis by a secreted lipocalin that is transcriptionally regulated by il-3 deprivation. Science *293*, 829-34.

Dougherty, D. A. (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science *271*, 163-8.

Ellison, R. T. (1994). The effects of lactoferrin on gram-negative bacteria. Adv Exp Med Biol *357*, 71-90.

Fantone, J. C., and Ward, P. A. (1985). Polymorphonuclear leukocyte-mediated cell and tissue injury: oxygen metabolites and their relations to human disease. Hum Pathol *16*, 973-8.

Ferguson, A. D., Hofmann, E., Coulton, J. W., Diederichs, K., and Welte, W. (1998). Siderophore-mediated iron transport: crystal structure of FhuA with bound lipopolysaccharide. Science *282*, 2215-20.

Fernandez-Pol, J. A. (1978). Isolation and characterization of a siderophore-like growth factor from mutants of SV40-transformed cells adapted to picolinic acid. Cell *14*, 489-99.

Flower, D. R. (1996). The lipocalin protein family: structure and function. Biochem J *318*, 1-14.

Flower, D. R. (1995). Multiple molecular recognition properties of the lipocalin protein family. J Mol Recognit *8*, 185-95.

Flower, D. R., North, A. C., and Attwood, T. K. (1991). Mouse oncogene protein 24p3 is a member of the lipocalin protein family. Biochem Biophys Res Commun *180*, 69-74.

Flower, D. R., North, A. C., and Attwood, T. K. (1993). Structure and sequence relationships in the lipocalins and related proteins. Protein Sci *2*, 753-61.

Flower, D. R., North, A. C., and Sansom, C. E. (2000). The lipocalin protein family: structural and sequence overview. Biochim Biophys Acta *1482*, 9-24.

Friedl, A., Stoesz, S. P., Buckley, P., and Gould, M. N. (1999). Neutrophil gelatinase-associated lipocalin in normal and neoplastic human tissues. Cell type-specific pattern of expression. Histochem J *31*, 433-41.

Gallivan, J. P., and Dougherty, D. A. (1999). Cation-pi interactions in structural biology. Proc Natl Acad Sci U S A *96*, 9459-64.

Ganz, T., Selsted, M. E., Szklarek, D., Harwig, S. S., Daher, K., Bainton, D. F., and Lehrer, R. I. (1985). Defensins. Natural peptide antibiotics of human neutrophils. J Clin Invest *76*, 1427-35.

Garay-Rojas, E., Harper, M., Hraba-Renevey, S., and Kress, M. (1996). An apparent autocrine mechanism amplifies the dexamethasone- and retinoic acid-induced expression of mouse lipocalin-encoding gene 24p3. Gene *170*, 173-80.

Gibson, F., and Magrath, D. I. (1969). The isolation and characterization of a hydroxamic acid (aerobactin) formed by Aerobacter aerogenes 62-I. Biochim Biophys Acta *192*, 175-84.

Gouet, P., Courcelle, E., Stuart, D. I., and Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics *15*, 305-8.

Haddy, T. B., Rana, S. R., and Castro, O. (1999). Benign ethnic neutropenia: what is a normal absolute neutrophil count? J Lab Clin Med *133*, 15-22.

Harrison, P. M., and Arosio, P. (1996). The ferritins: molecular properties, iron storage function and cellular regulation. Biochim Biophys Acta *1275*, 161-203.

Harvath, L. (1991). Neutrophil chemotactic factors. Exs *59*, 35-52.

Hershko, C., Graham, G., Bates, G. W., and Rachmilewitz, E. A. (1978). Non-specific serum iron in thalassaemia: an abnormal serum iron fraction of potential toxicity. Br J Haematol *40*, 255-63.

Hollifield, W. C., Jr., and Neilands, J. B. (1978). Ferric enterobactin transport system in Escherichia coli K-12. Extraction, assay, and specificity of the outer membrane receptor. Biochemistry *17*, 1922-8.

Holm, L., and Sander, C. (1998). Touring protein fold space with Dali/FSSP. Nucleic Acids Res *26*, 316-9.

Hraba-Renevey, S., Turler, H., Kress, M., Salomon, C., and Weil, R. (1989). SV40-induced expression of mouse gene 24p3 involves a post-transcriptional mechanism. Oncogene *4*, 601-8.

Hubbard, S. J., and Thornton, J. M. (1993). NACCESS (London: Department of Biochemistry and Molecular Biology, University College).

Jones, R. L., Peterson, C. M., Grady, R. W., and Cerami, A. (1980). Low molecular weight iron-binding factor from mammalian tissue that potentiates bacterial growth. J Exp Med *151*, 418-28.

Jr., C. W. C., and Sweet, R. M. (1997). Macromolecular Crystallography Part A. In Methods in Enzymology, J. N. Abelson and M. I. Simon, eds. (San Diego: Academic Press).

Jr., C. W. C., and Sweet, R. M. (1997). Macromolecular Crystallography Part B. In Methods in Enzymology, J. N. Abelson and M. I. Simon, eds. (San Diego: Academic Press).

Karpishin, T. B., Dewey, T. M., and Raymond, K. N. (1993). The Vanadium (IV) Enterobactin Complex: Structural, Spectroscopic, and Electrochemical Characterization. J Am Chem Soc *115*, 1842-1851.

Kasik, J. W., and Rice, E. J. (1995). An increase in expression of the lipocalin 24p3 is found in mouse uterus coincident with birth. Am J Obstet Gynecol *173*, 613-7.

Keller, H. U. (1983). Motility, cell shape, and locomotion of neutrophil granulocytes. Cell Motil *3*, 47-60.

Keller, H. U. (1983). Shape, motility and locomotor responses of neutrophil granulocytes. Agents Actions Suppl *12*, 54-72.

Kissenger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999). Rapid automated molecular replacement by evolutionary search. Acta Cryst *D55*, 484-491.

Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999). Rapid automated molecular replacement by evolutionary search. Acta Crystallogr D Biol Crystallogr *55*, 484-91.

Kjeldsen, L., Johnsen, A. H., Sengelov, H., and Borregaard, N. (1993). Isolation and primary structure of NGAL, a novel protein associated with human neutrophil gelatinase. J Biol Chem *268*, 10425-32.

Klebanoff, S. J. (1980). Oxygen metabolism and the toxic properties of phagocytes. Ann Intern Med *93*, 480-9.

Klebba, P. E., McIntosh, M. A., and Neilands, J. B. (1982). Kinetics of biosynthesis of iron-regulated membrane proteins in Escherichia coli. J Bacteriol *149*, 880-8.

Kleywegt, G. J., and Brunger, A. T. (1996). Checking your imagination: applications of the free R value. Structure *4*, 897-904.

Kolkenbrock, H., Hecker-Kia, A., Orgel, D., Kinawi, A., and Ulbrich, N. (1996). Progelatinase B forms from human neutrophils. complex formation of monomer/lipocalin with TIMP-1. Biol Chem *377*, 529-33.

Kuzmic, P. (1996). Program DYNAFIT for the Analysis of Enzyme Kinetic Data: Application to HIV Proteinase. Anal. Biochem. *237*, 260-273.

Lee, C.-W., Ecker, D. J., and Raymond, K. N. (1985). The pH-Dependant Reduction of Ferric Enterobactin Probed by Electrochemical Methods and Its Implications for Microbial Iron Transport. J Am Chem Soc *107*, 6920-6923.

Liu, Q., Ryon, J., and Nilsen-Hamilton, M. (1997). Uterocalin: a mouse acute phase protein expressed in the uterus around birth. Mol Reprod Dev *46*, 507-14.

Locher, K. P., Rees, B., Koebnik, R., Mitschler, A., Moulinier, L., Rosenbusch, J. P., and Moras, D. (1998). Transmembrane signaling across the ligand-gated FhuA receptor: crystal structures of free and ferrichrome-bound states reveal allosteric changes. Cell *95*, 771-8.

Locher, K. P., and Rosenbusch, J. P. (1997). Oligomeric states and siderophore binding of the ligand-gated FhuA protein that forms channels across Escherichia coli outer membranes. Eur J Biochem *247*, 770-5.

Loomis, L. D., and Raymond, K. N. (1991). Solution Equilibria of Enterobactin and Metal-Enterobactin Complexes. Inorganic Chemistry *30*, 906-911.

Mary, J. Y. (1984). Normal human granulopoiesis revisited. I. Blood data. Biomed Pharmacother *38*, 33-43.

McRee, D. E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. J Struct Biol *125*, 156-65.

Meheus, L. A., Fransen, L. M., Raymackers, J. G., Blockx, H. A., Van Beeumen, J. J., Van Bun, S. M., and Van de Voorde, A. (1993). Identification by microsequencing of lipopolysaccharide-induced proteins secreted by mouse macrophages. J Immunol *151*, 1535-47.

Mollinedo, F., Borregaard, N., and Boxer, L. A. (1999). Novel trends in neutrophil structure, function and development. Immunol Today *20*, 535-7.

Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992). Stereochemical quality of protein structure coordinates. Proteins *12*, 345-364.

Neilands, J. B. (1995). Siderophores: structure and function of microbial iron transport compounds. J Biol Chem *270*, 26723-6.

Newton, S. M., Igo, J. D., Scott, D. C., and Klebba, P. E. (1999). Effect of loop deletions on the binding and transport of ferric enterobactin by FepA. Mol Microbiol *32*, 1153-65.

Nielsen, B. S., Borregaard, N., Bundgaard, J. R., Timshel, S., Sehested, M., and Kjeldsen, L. (1996). Induction of NGAL synthesis in epithelial cells of human colorectal neoplasia and inflammatory bowel diseases. Gut *38*, 414-20.

O'Brien, I. G., and Gibson, F. (1970). The structure of enterochelin and related 2,3-dihydroxy-N-benzoylserine conjugates from Escherichia coli. Biochim Biophys Acta *215*, 393-402.

Otto, B. R., Verweij-van Vught, A. M., and MacLaren, D. M. (1992). Transferrins and heme-compounds as iron sources for pathogenic bacteria. Crit Rev Microbiol *18*, 217-33.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray Diffraction Data Collected in Oscillation Mode. In Methods in Enzymology, C. W. C. Jr. and R. M. Sweet, eds. (NY: Academic Press), pp. 307-326.

Picker, L. J. (1994). Control of lymphocyte homing. Curr Opin Immunol *6*, 394-406.

Pollack, J. R., and Neilands, J. B. (1970). Enterobactin, an iron transport compound from Salmonella typhimurium. Biochem Biophys Res Commun *38*, 989-92.

Ponka, P., Beaumont, C., and Richardson, D. R. (1998). Function and regulation of transferrin and ferritin. Semin Hematol *35*, 35-54.

Ratledge, C., and Dover, L. G. (2000). Iron metabolism in pathogenic bacteria. Annu Rev Microbiol *54*, 881-941.

Reed, R. J. (1986). Improved Fourier coefficients for maps using phases from partial structures with errors. Acta Cryst *A42*, 140-149.

Roberts, R. B., Abelson, P. H., Cowie, D. B., Bolton, E. T., and Britten, R. J. (1963). Studies of biosynthesis in *Escherichia coli*. In Publication 607 (Washington D.C.: Carnegie Institue of Washington).

Schomaker, V., and Trueblood, K. N. (1998). Acta Cryst *B54*, 507-514.

Sengelov, H., Boulay, F., Kjeldsen, L., and Borregaard, N. (1994). Subcellular localization and translocation of the receptor for N-formylmethionyl-leucyl-phenylalanine in human neutrophils. Biochem J *299*, 473-9.

Stewart, J. M. (2000). The cytoplasmic fatty-acid-binding proteins: thirty years and counting. Cell Mol Life Sci *57*, 1345-59.

Stoesz, S. P., and Gould, M. N. (1995). Overexpression of neu-related lipocalin (NRL) in neu-initiated but not ras or chemically initiated rat mammary carcinomas. Oncogene *11*, 2233-41.

Strong, R. K., Bratt, T., Cowland, J. B., Borregaard, N., Wiberg, F. C., and Ewald, A. J. (1998). Expression, purification, crystallization and crystallographic characterization of dimeric and monomeric human neutrophil gelatinase associated lipocalin (NGAL). Acta Crystallogr D Biol Crystallogr *54*, 93-5.

Sullivan, S. J., Daukas, G., and Zigmond, S. H. (1984). Asymmetric distribution of the chemotactic peptide receptor on polymorphonuclear leukocytes. J Cell Biol *99*, 1461-7.

Thomas, E. L., Lehrer, R. I., and Rest, R. F. (1988). Human neutrophil antimicrobial activity. Rev Infect Dis *10 Suppl 2*, S450-6.

Triebel, S., Blaser, J., Reinke, H., and Tschesche, H. (1992). A 25 kDa alpha 2-microglobulin-related protein is a component of the 125 kDa form of human gelatinase. FEBS Lett *314*, 386-8.

Waksman, G. (1994). Crystal structure of the phosphotyrosine recognition domain SH2 of the Src oncogene product complexed with tyrosine-phosphorylated peptides. Cell Mol Biol (Noisy-le-grand) *40*, 611-8.

Walker, R. I., and Willemze, R. (1980). Neutrophil kinetics and the regulation of granulopoiesis. Rev Infect Dis *2*, 282-92.

Wang-Iverson, P., Pryzwansky, K. B., Spitznagel, J. K., and Cooney, M. H. (1978). Bactericidal capacity of phorbol myristate acetate-treated human polymorphonuclear leukocytes. Infect Immun *22*, 945-55.

Wavefunction. (1992). SPARTAN (Ivine, CA.

Weinberg, E. D. (1984). Iron withholding: a defense against infection and neoplasia. Physiol Rev *64*, 65-102.

Whited, S. C., and Gallin, J. I. (1980). Neutrophil chemotaxis. Int J Dermatol *19*, 130-8.

Winn, M. D., Isupov, M. N., and Murshudov, G. N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. Acta Crystallogr D Biol Crystallogr *57*, 122-33.

Wookey, P., and Rosenberg, H. (1978). Involvement of inner and outer membrane components in the transport of iron and in colicin B action in Escherichia coli. J Bacteriol *133*, 661-6.

Wooldridge, K. G., and Williams, P. H. (1993). Iron uptake mechanisms of pathogenic bacteria. FEMS Microbiol Rev *12*, 325-48.

Xu, S. Y., Carlson, M., Engstrom, A., Garcia, R., Peterson, C. G., and Venge, P. (1994). Purification and characterization of a human neutrophil lipocalin (HNL) from the secondary granules of human neutrophils. Scand J Clin Lab Invest *54*, 365-76.

Zerega, B., Cermelli, S., Michelis, B., Cancedda, R., and Cancedda, F. D. (2000). Expression of NRL/NGAL (neu-related lipocalin/neutrophil gelatinase-associated lipocalin) during mammalian embryonic development and in inflammation. Eur J Cell Biol *79*, 165-72.

Vita

David Goetz was born in Albuquerque, New Mexico. He has lived in many places throughout the west and southwest. Currently, he lives in Seattle. In 1997, he earned his Bachelor of Arts degree in Biochemistry from Rice University in Houston, Texas. In 2002, he earned a Doctor of Philosophy at the University of Washington in Molecular and Cellular Biology.