Ligand Preference Inferred from the Structure of Neutrophil Gelatinase Associated Lipocalin †

ARTICLE *in* BIOCHEMISTRY · MARCH 2000

Impact Factor: 3.02 · DOI: 10.1021/bi992215v · Source: PubMed

6 AUTHORS, INCLUDING:

Roland K Strong
Fred Hutchinson Cancer Research Center
107 PUBLICATIONS **6,699** CITATIONS

Ligand Preference Inferred from the Structure of Neutrophil Gelatinase Associated Lipocalin†

David H. Goetz,‡ Sirkku T. Willie,‡ Roger S. Armen,‡ Tomas Bratt,§ Niels Borregaard,§ and Roland K. Strong*,‡

Division of Basic Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, and The Granulocyte Research Laboratory, Department of Hematology, Rigshospitalet, Copenhagen, Denmark

Received September 23, 1999; Revised Manuscript Received November 29, 1999

ABSTRACT: Neutrophil gelatinase associated lipocalin (NGAL), a constituent of neutrophil granules, is a member of the lipocalin family of binding proteins. NGAL can also be highly induced in epithelial cells in both inflammatory and neoplastic colorectal disease. NGAL is proposed to mediate inflammatory responses by sequestering neutrophil chemoattractants, particularly N-formylated tripeptides and possibly leukotriene B_4 and platelet activating factor. The crystal structures of NGAL display a typical lipocalin fold, albeit with an unusually large and atypically polar binding site, or calyx. The fold of NGAL is most similar to the epididymal retinoic acid-binding protein, another lipocalin, though the overall architecture of the calyces are very different. The crystal structures also reveal either sulfate ions or an adventitiously copurified fatty acid bound in the binding site. Neither ligand is displaced by added N-formylated tripeptides. The size, shape, and character of the NGAL calyx, as well as the low relative affinity for N-formylated tripeptides, suggest that neither the copurified fatty acid nor any of the proposed ligands are likely to be the preferred ligand of this protein. Comparisons between the crystal structures and the recently reported solution structure of NGAL reveal significant differences, in terms of both the details of the structure and the overall flexibility of the fold.

Various mediators of inflammation, such as prostaglandins, platelet activating factor (PAF),[1] leukotrienes, interleukin 8 (IL-8), and granulocyte colony stimulating factor, are chemotactic for neutrophils, which are recruited to inflammatory foci, resulting in the augmentation and progression of the immune response (*1*). Activated neutrophils release a variety of membrane-associated proteins and soluble compounds from intracellular granules that mediate neutrophil functions: endothelial adhesion, migration across basement membranes, and the phagocytosis and destruction of pathogenic microorganisms (*2*). *N*eutrophil *g*elatinase *a*ssociated *l*ipocalin (NGAL), a member of the lipocalin family, is released from neutrophil granules as a 25 kDa monomer, a 46 kDa disulfide-linked homodimer, and a disulfide-linked heterodimer with gelatinase B (matrix metalloproteinase 9) (*3*, *4*). NGAL is expressed in immature neutrophil precursors (*5*, *6*) and in epithelial cells during both inflammation and neoplastic transformation (*1*). Both the murine (24p3) and the rat analogue of NGAL (*neu*-related lipocalin) were originally identified in screens for genes overexpressed during tumorigenesis (*7*, *8*).

Lipocalins are a diverse family of small proteins that generally bind small, hydrophobic ligands but which can also bind soluble, extracellular macromolecules and specific cell-surface receptors (*9*). Lipocalins are involved in many processes: retinol transport, invertebrate cryptic coloration, olfaction, pheromone transport, and prostaglandin synthesis. Despite limited sequence similarity, the lipocalin fold is remarkably well conserved. The core structure consists of an eight-stranded, antiparallel, continuously hydrogen-bonded β-barrel which defines a calyx- or cup-shaped structure that encloses the ligand binding site (*9*). The strand connectivity in the barrel is consistently +1. The loops linking the strands are typically short β-hairpins, except the first loop, which describes a large Ω loop that usually folds back onto the barrel, partially constricting the binding site. There is a short 3_{10} helix at the N-terminus and an α-helix at the C-terminus.

Unlike other lipocalins, NGAL shows no affinity for retinoic acid (REA) (*10*) but is reported to bind the tripeptide *N*-formyl-Met-Leu-Phe (fMLF), a potent neutrophil chemoattractant (*11*), and possibly other lipophilic mediators of inflammation such as PAF, leukotriene B_4 (LTB_4), and LPS (*1*). Through these interactions, NGAL is proposed to have important immunomodulatory functions. The role of NGAL in tissue remodeling and tumorigenesis is not known. We report the crystal structure of two forms of NGAL in order to define its interactions with proposed ligands and thereby delimit its role in mediating immune responses. The solution structure of NGAL has recently been reported (*12*).

† This work was supported by a grant from the National Institutes of Health to R.K.S. (AI42200) and a National Science Foundation Graduate Research Fellowship to D.H.G.

* To whom correspondence should be addressed. E-mail: rstrong@fhcrc.org.

‡ Fred Hutchinson Cancer Research Center.

§ Rigshospitalet.

[1] Abbreviations: βLG, β-lactoglobulin; ERBP, epididymal retinoic acid-binding protein; fMLF, *N*-formyl-Met-Leu-Phe; IL-8, interleukin 8; LTB_4, leukotriene B_4; MUP, mouse major urinary protein; NCA, *n*-capric acid; NGAL, neutrophil gelatinase associated lipocalin; NP, nitrophorin; PAF, platelet activating factor; REA, retinoic acid.

10.1021/bi992215v CCC: $19.00
Published on Web 02/02/2000

EXPERIMENTAL PROCEDURES

Crystals of baculovirus-expressed, dimeric NGAL were grown by vapor diffusion at 18 °C from solutions containing 20 mg mL^{-1} protein, 18−22% w/w poly(ethylene glycol) (average M_r = 8000), 15% v/v glycerol, and 100 mM sodium acetate (pH = 4.5) as previously described (*13*). The space group of the dimer crystals is $P4_12_12$, $a = b = 54.29$ Å, and $c = 122.0$ Å, with half a dimer in the asymmetric unit (solvent content ≈30%). The monomer form of NGAL was crystallized by vapor diffusion at 18 °C from solutions containing 10 mg mL^{-1} protein, 21−23% w/w poly(ethylene glycol) (average M_r = 8000), 12−14% saturated ammonium sulfate, 10% v/v glycerol, 25 mM PIPES (pH = 7.0), and 10 mM dithiothreitol to reduce trace contaminating dimer. The space group of the monomer crystals is $P4_12_12$, $a = b = 115.4$ Å, and $c = 118.0$ Å, with two molecules in the asymmetric unit (solvent content ≈71%). Data were collected from cryopreserved crystals (*14*) of dimeric NGAL on a Rigaku R-AXIS 4c image plate detector and monomeric NGAL at beamline 5.0.2 at the Advanced Light Source (Berkeley, CA). Crystals were grown with protein purified at pH = 7.0 mixed with the proposed tripeptide ligand, fMLF (Sigma), at a peptide:protein molar ratio of 3:1. Stock solutions were prepared by dissolving the peptide in dimethylformamide at 20 mg mL^{-1}. A single heavy atom derivative was obtained by soaking dimer crystals in mother liquor containing 1 mM di-*m*-iodobis(ethylenediamine)-diplatinum(II) nitrate (PIP) for 24 h. The data were processed and reduced with DENZO and SCALEPACK (*15*) (see Table 1). Difference Patterson syntheses were inspected with XtalView (*16*). Heavy atom parameters were refined with SHARP (*17*); phases were extended from 3.91 to 2.5 Å using the SOLOMON (*18*) solvent-flattening routine as implemented in SHARP with an assigned solvent content of 49.5%. An initial electron density map for the dimer calculated with these phases was readily interpretable. A partially refined model of NGAL was used as a search model in molecular replacement phasing of the monomer data using EPMR (*19*). Two solutions were readily found (correlation coefficient = 0.674; *R*-factor = 35.8% after rigid body refinement in EPMR). Models were built with Xfit (*16*) and O (*20*) and refined initially with X-PLOR (*21*) and subsequently with CNS (*22*) using the maximum likelihood target function mlf. A bulk solvent correction and *B*-factor refinement were applied in the later stages of refinements. Intermediate structures were manually compared against simulated annealing omit and SIGMAA (*23*)-weighted $2F_o - F_c$ and $F_o - F_c$ electron density maps. The progress of the refinements was confirmed by the monotonic decrease in both R_{cryst} and R_{free} (*24*). NCS restraints were not imposed in the refinement of the monomer structures in accordance with the R_{free}. Geometry was evaluated with CNS and PROCHECK (*25*). Residues 1−3 in the dimer and monomer B are disordered, as are residues 1−4 in monomer A. Additional residues have been left as alanines due to the poor quality of the local density: Thr-4, Glu-61, Lys-62, Arg-72, Lys-73, Lys-74, and Lys-98 in the dimer and Thr-4 in monomer B. Late in the refinement process, ligands were built into density not accounted for by protein, carbohydrate, or ordered solvent: an NCA molecule in the dimer structure and sulfate ions in the monomer structures. An analysis of the nonbonded atomic interactions [using ERRAT (*26*)] shows that 98.9%, 97.7%, or 99.4% of the residues are within the 95% confidence levels for the dimer, monomer A, and monomer B structures, respectively; no residues are outside of the 99% confidence levels. Five of the seven residues outside of the 95% confidence level are involved in crystal contacts or are at an N-linked oligosaccharide attachment site. The average per residue 3D profile scores for the three models, calculated with Verify3D (*27*), are all 0.46. Coordinates have been deposited with the Research Collaboratory for Structural Bioinformatics (accession numbers 1QQS and 1DFV).

Table 1: Data Collection and Refinement Statistics

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 (%)[a]	88.1 (38.7)	90.2 (49.7)	87.9 (60.0)
$\langle I/\sigma(I)\rangle$	20.4 (5.2)	17.1 (2.2)	43.8 (7.6)
R_{sym} (%)[b]	5.2 (10.9)	5.6 (10.9)	5.2 (25.5)
no. of sites		4	
isomorphous difference (%)		18.6	
Kraut scale (%)		96.5	
% anomalous data		58.3	
empirical *K*		5.56	
phasing power (centrics)[c]		2.28	
phasing power (acentrics)[c]		2.73	
figure of merit (to 3.91 Å; %)		64.6	
refinement			
resolution (Å)	20−2.4		20−2.6
reflections (all $F > 0$)	6110		21861
protein atoms	1392		2821
solvent atoms	74		133
CHO atoms	39		42
ligand atoms	12		10
R_{cryst} (%)[d]	21.6		28.1
R_{free} (%)[d]	28.5		29.3
geometry (rmsd from ideality)			
bond length (Å)	0.008		0.008
bond angles (deg)	1.4		1.7
dihedrals (deg)	27.0		28.1
average *B* (Å^2)	24.5		46.8
average protein *B* (Å^2)	23.9		50.3 (A)/ 43.5 (B)
average helix *B* (Å^2)	27.0		53.0 (A)/ 42.1 (B)
Ramachandran			
most favored (%)	77		80.7
disallowed (%)	0		0

[a] Values for the highest resolution shell are shown in parentheses. [b] $R_{sym} \equiv \sum|I - \langle I\rangle|/\sum\langle I\rangle$, where I is the observed intensity and $\langle I\rangle$ is the mean intensity of multiple observations of symmetry-related reflections. [c] Phasing power $\equiv \langle|F_H|\rangle/E$, where F_H is the heavy atom structure factor amplitude and E is the residual lack-of-closure error. [d] R_{cryst}, $R_{free} \equiv \sum||F_o| - |F_c||/\sum|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively. R_{free} is calculated for a randomly chosen 10% of the reflections excluded from refinement (*24*).

RESULTS

The crystal structure of dimeric human NGAL, crystallized at an acidic pH, was determined at 2.5 Å resolution by isomorphous replacement; the crystal structure of monomeric human NGAL, crystallized at neutral pH, was determined at 2.6 Å resolution by molecular replacement using the dimer structure as a search model (Table 1 and Figure 1). There is half of a dimer in the asymmetric unit of the dimer crystals

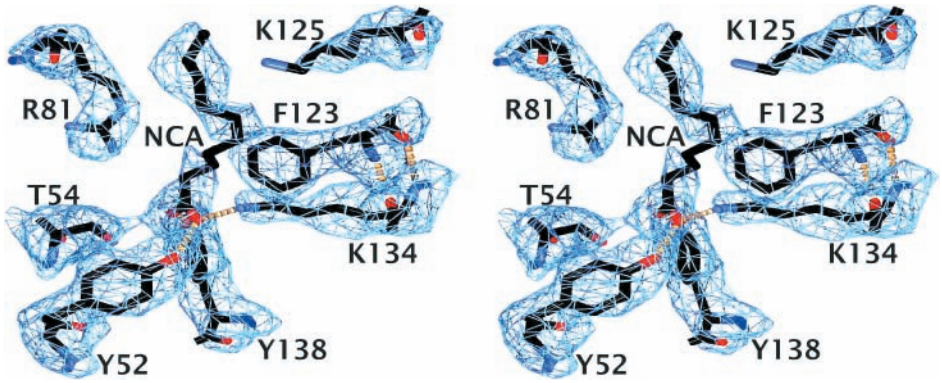

FIGURE 1: Electron density around the observed fatty acid ligand. Stereoview of the simulated annealing, $F_o - F_c$ omit maps (blue net) in the vicinity of the bound fatty acid (NCA) in the pH = 4.5 dimer structure, contoured at 3σ, with all neighboring residues shown. All atoms shown were omitted in the calculations and are colored as follows: carbon, black; oxygen, red; nitrogen, blue. The NCA density is connected and continuous at lower contouring levels. All figures were generated with SwissPDBViewer and rendered with POV-RAY3 except where noted.

and two monomers in the asymmetric unit of the monomer crystals, yielding three independent views of the NGAL structure. Despite large differences in pH and salt concentration, resulting in very different arrangements of ordered solvent, the three NGAL structures are essentially identical and display a typical lipocalin fold, with the conserved eight-stranded β-barrel, a short 3_{10} helix (residues 24−28) that packs against a conserved tryptophan (*31*) at Phe-27, and a four-turn α-helix (residues 145−160; Figure 2A). A ninth, three-residue β-strand (residues 165−167) packs against the first β-stand, completing the lipocalin fold. However, the N-terminus of the NGAL structures is more ordered than in most lipocalin structures and folds back against the side of the barrel passing over the sixth, seventh, and eighth β-strands (Figure 2A). Another two-turn 3_{10} helix is present within this arm (residues 12−16). Loop 1, rather than packing into the calyx as in most lipocalins, lays back against the ninth β-strand.

The largest difference in the three NGAL structures occurs at the N-termini, where the first two ordered residues (4 and 5) have different structures, affected by different crystal packing environments. The largest backbone change in the remainder of the structure is at loop 3, which moves by approximately 3 Å at Pro-101 from the dimer to the monomer structures. Within the calyx, only residues Leu-70, Trp-79, and Arg-81 show significantly altered conformations among the NGAL structures. Leu-70 occupies three different rotomers in the three different structures, the side chain of Trp-79 is disordered in the dimer structure but not the two monomer structures, and the side chain of Arg-81 either points toward the bottom of the calyx in the dimer structure or points upward in the monomer structures. There is a disulfide linkage between cysteines 76 and 175 in all three structures. Formation of the disulfide bond between Cys-87 in the dimer does not occlude or affect the structure of the calyx. Cys-87 lies in hairpin turn between the fourth and fifth β-strands that likely allows some interdomain flexibility in the homodimer or the gelatinase B heterodimer. Other than the disulfide bond there is no other interaction between the two molecules in the homodimer. There are three ordered N-linked carbohydrates at Asn-65 in the dimer structure, two in monomer A and one in monomer B. The lack of a dramatic pH-induced conformational change in NGAL contrasts with β-lactoglobulin (βLG), where a pH change from 6.2 to 8.2 results in an approximately 10 Å movement of the loop connecting the fifth and sixth β-strands (loop 3, Figure 2) at Leu-87, occluding the binding site at low pH (*28*).

The NMR solution structure of NGAL (*12*) differs considerably from the crystal structures (rms deviations of 4.45, 4.48, and 4.46 Å on all Cαs between the regularized mean NMR structure and the half-dimer, monomer A, and monomer B respectively). The N-terminal arm, starting at 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 rms deviations of 2.99, 2.98, and 3.00 Å (on Cαs) between the regularized mean solution structure and the half-dimer, monomer A, and monomer B, respectively. The average rms deviation between the 20 convergent NMR structures (2.18 Å on all backbone atoms) is larger than the deviations between the three independent NGAL crystal structures by more than 3-fold (Figure 2A). Analysis of the NMR results also suggests that the α-helix may represent a rigid structural element flexibly linked to the rest of the protein. However, in the crystal structures, the α-helices adopt identical conformations, with thermal parameters indistinguishable from the protein as a whole (Table 1). These deviations are unlikely to reflect differences between solution and crystal environments since radically different crystallization conditions yield essentially identical structures, suggesting that NGAL has a fairly rigid structure.

A search of the structural database reveals that the structure of NGAL is most closely related to the structures of the epididymal retinoic acid-binding protein (*29, 30*) (ERBP; Figure 2B) and the mouse major urinary protein (*31*) (MUP; Figure 2C) and somewhat less closely related to the structures of βLG (*32*) and nitrophorin (*33*) (NP). While the folds of NGAL and ERBP are nearly indistinguishable, loops 1 and 3 in MUP pack well into and over the calyx, resulting in a small, sealed binding cavity (Figures 2C and 3C). The calyx of NGAL is unusually large and open compared to other lipocalins (Figures2 and 3) but shallower than the binding cavities of ERBP, βLG, or NP. The NGAL calyx is considerably larger even than that of NP, which binds a substituted heme (Figure 4). In addition to a much larger enclosed volume, the NGAL calyx is also uncharacteristically lined with many polar and positively charged residues.

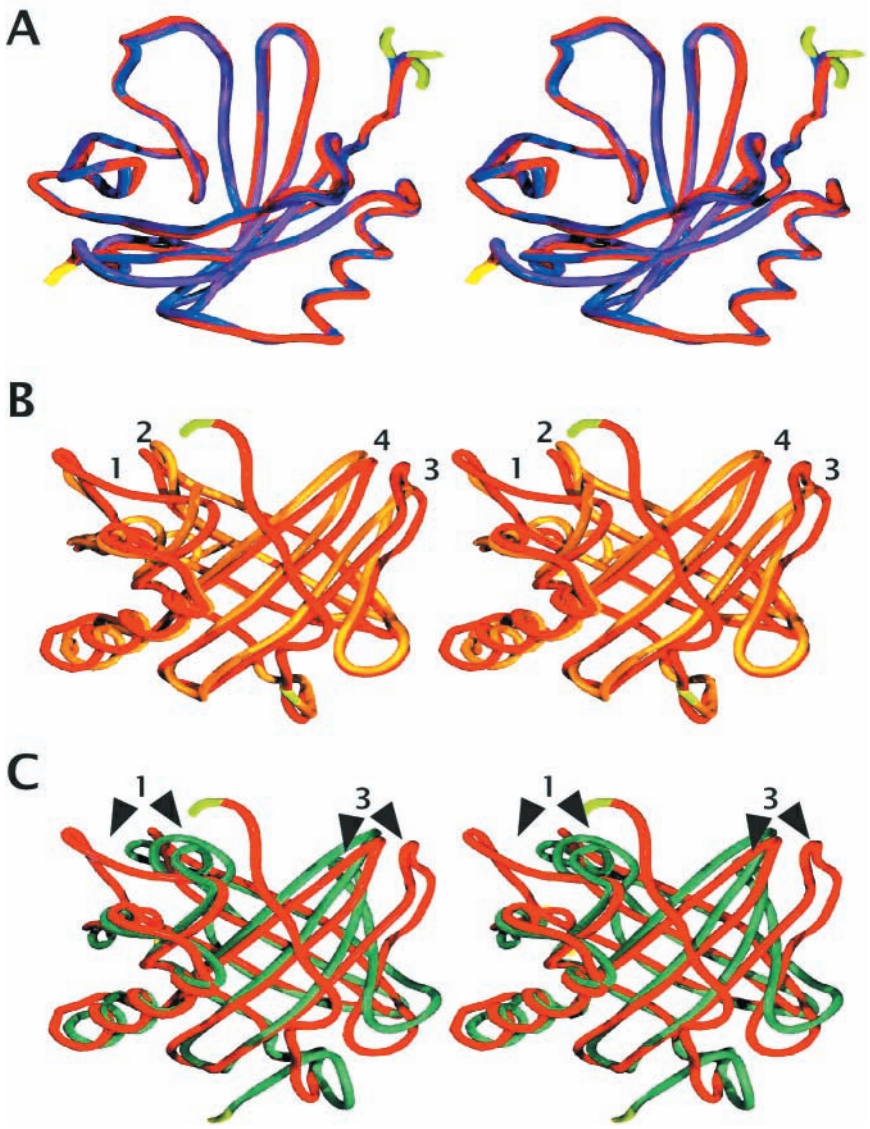

FIGURE 2: Structural comparisons of NGAL and other lipocalins. (A) Stereoview of the superposition of the three independently determined NGAL structures, shown as backbone ribbons, colored as follows: half-dimer, red; monomer A, blue; monomer B, purple; N-termini, green; C-termini, yellow. rms deviations for the superpositions are monomer A/monomer B = 0.24 Å (on all Cαs), 0.35 Å (on all backbone atoms); monomer A/half-dimer = 0.65 Å (Cαs), 0.71 Å (backbone); and monomer B/half-dimer = 0.63 Å (Cαs), 0.65 Å (backbone). Stereoview of the superposition of the NGAL half-dimer (pH = 4.5; shown in red) on (B) the epididymal retinoic acid-binding protein (ERBP; shown in orange) and (C) the murine major urinary protein (MUP; shown in green) shown as backbone ribbons. Structures were superimposed using the DALI server (*36*). rms deviations for superpositions are NGAL and ERBP, 23% identical, 1.9 Å (on 145 Cαs); NGAL and MUP, 16% identical, 2.0 Å (on 141 Cαs); NGAL and βLG, 15% identical, 2.5 Å (on 139 Cαs); and NGAL and NP, 10% identical, 3.4 Å (on 154 Cαs). N-Termini are colored green; C-termini, yellow. Loops that define the upper end of the binding pockets are numbered. Loop 1 consists of residues 38−50 in NGAL, residues 21−35 in ERBP (this loop is disordered in the ERBP structures in the presence or absence of ligand), and residues 30−44 in MUP; loop 2, residues 69−81 in NGAL, residues 53−65 in ERBP, and residues 61−73 in MUP; loop 3, residues 95−105 in NGAL, residues 78−84 in ERBP, and residues 86−92 in MUP; loop 4, residues 125−132 in NGAL, residues 104−111 in ERBP, and residues 111−118 in MUP.

Late in the refinement process, electron density in difference Fourier syntheses was found in the NGAL calyx of all three structures that did not correspond to protein, ordered solvent, or carbohydrate (Figure 1). In the dimer structure, the density was consistent with a fatty acid with a nine carbon long tail, decanoic or *n*-capric acid; NCA). Mass spectroscopy studies to confirm the identification of the copurified ligand have been inconclusive (data not shown). It is possible that the fatty acid bound in this site has a longer aliphatic tail that is disordered and therefore not visible in the difference Fouriers or is a mixture of different fatty acids and/or related compounds.

The carboxylate of the NCA hydrogen bonds to the side chains of Tyr-52, Lys-134, and Tyr-138 and lies in a pocket lined by the side chains of Thr-54, Tyr-56, Arg-81, Phe-123, and Thr-136 (Figure 1). This pocket lies at the deepest point in the NGAL calyx (Figure 3A). REA sits approximately 8.5 Å deeper in the ERBP calyx than does NCA in NGAL. The side chains of three residues (Tyr-56, Phe-123, and Tyr-138) form the floor of the NGAL calyx at this point, closing off a cavity, approximately 780 $Å^3$ in volume, that roughly corresponds to the bottom end of the REA binding site in ERBP. This cavity in the structure of NGAL is lined with the side chains of Phe-22, Phe-27, Tyr-56, Tyr-

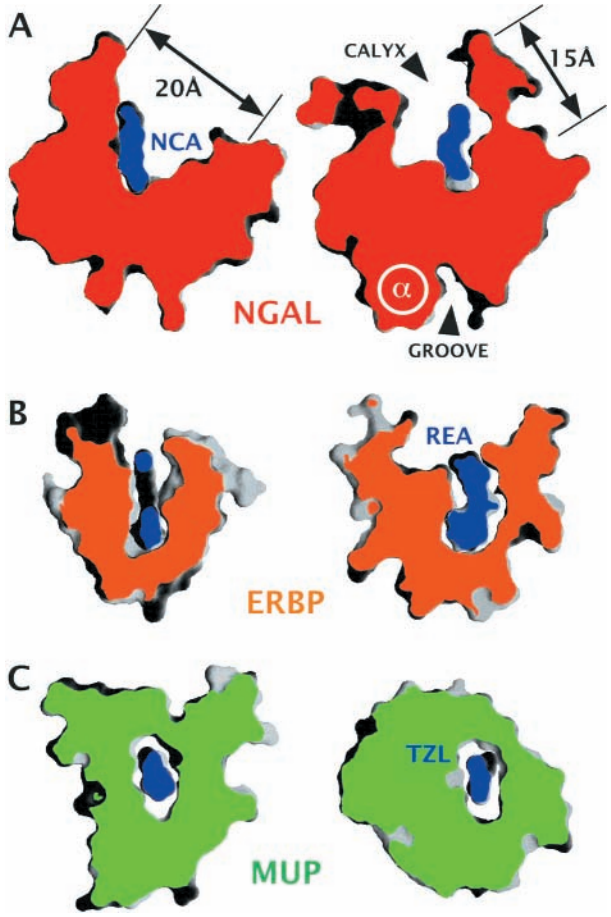

FIGURE 3: Orthogonal, 3 Å thick slices through GRASP (*37*) molecular surfaces of the NGAL half-dimer (A), colored red, ERBP (B), colored orange, and MUP (C), colored green. The ligands in each case have been colored blue and labeled *n*-capric acid (NCA) in NGAL, retinoic acid (REA) in ERBP, and 2-(*sec*-butyl)-thiazoline (TZL) in MUP. The approximate dimensions of the calyx, the location of the α-helix, and the groove have been indicated for NGAL. The groove is lined with the side chains of Ser-112, Gln-117, His-118, Met-120, Leu-137, Thr-141, Glu-143, Thr-145, Glu-147, Leu-148, Asn-151, and Phe-155. The molecules have been oriented so that the ligands are approximately vertical.

64, Phe-83, Phe-92, Val-110, and Val-121. Assignment of the carboxylate binding pocket is bolstered by its similarity to the carboxylate pocket of the palmitate binding site of the fatty acid binding protein of *Manduca sexta* (*34*). In this protein, the carboxylate of the palmitate hydrogen bonds to the side chains of Gln-39, Arg-127, and Tyr-129 and is enclosed by the side chains of Phe-103, Lys-105, Val-114, Ile-116, and a phosphate ion. The tail of the NCA molecule in NGAL lies in a shallow, apolar crevice defined by the side chains of Leu-94, Tyr-100, Leu-103, and Tyr-106 (Figure 4). However, the NCA does not come close to filling the NGAL calyx, nor is its shape and chemistry particularly complementary to the site it occupies, which contrasts many of the lipocalin ligands (Figures 3 and 4).

We presume that this bound fatty acid represents a serendipitous ligand copurified with NGAL. Since the recombinant protein used in the crystallization experiments is expressed in a baculovirus system, the ligand could represent a component of either insect cells or insect cell media, which contains fetal calf serum. The density is not consistent with any component of the crystallization mother liquor. During purification, the protein is exposed alternately to both low and neutral pH, implying that binding of the fatty acid is not pH dependent.

In both monomers, an ordered sulfate ion hydrogen bonds to the side chains of Arg-81, Tyr-106, Lys-125, and Lys-134, lying near the side chain of Phe-123. The position of the sulfates overlies the NCA molecule in the middle of the aliphatic tail when the three structures are superimposed. We interpret this result to mean that sulfate displaces the copurified fatty acid at the monomer crystallization conditions, which include approximately half molar ammonium sulfate. It is unlikely that pH plays a role in the differential binding of fatty acid versus sulfate, as the protein used in all the crystallization trials presumably retains bound NCA during exposure to a wide pH range. The differential binding of sulfate or NCA likely accounts for the observed change in the conformation of Arg-81, which points downward toward the NCA carboxylate in the dimer structures and points upward toward the sulfates in the monomer structures, and may indirectly affect the conformations of the side chains of Leu-70 and Trp-79. No other obvious ligand-induced conformational changes are observed in the NGAL structures.

In addition to the calyx, a small groove (approximately $10 \times 4 \times 4$ Å) lies between the α-helix and the sixth and seventh β-strands on the opposite side of NGAL (Figure 3A). A similar pocket was noted in the structure of βLG (*32*). Two of the heavy atom derivative sites lie in this groove (one site lies between His-118, Thr-145, and Leu148 and the other is at Met-120) so this structure may define an additional ligand binding site of unknown specificity.

The appearance of the density assigned as either NCA or sulfate in the difference Fourier syntheses was not altered when fMLF, one of the proposed ligands, was excluded from the crystallization conditions (data not shown), indicating that the observed density is unlikely to be a partially ordered fMLF molecule. Aside from the inconsistent electron density, fMLF sterically clashes with protein when docked into the NCA site. No additional electron density was observed in any other site on the protein, including the groove. These results suggest that, under the crystallization conditions, fMLF does not displace either NCA or sulfate bound in their respective pockets, nor does it bind to any other site on the surface of NGAL.

NGAL interacts with gelatinase B through a disulfide bond between Cys-87 (NGAL) and Cys-449 (the only free cysteine in gelatinase B). Cys-449 lies in a loop connecting the matrixin and hemopexin domains of the gelatinases. The recently reported crystal structure of gelatinase A (*35*), which has a domain organization similar to that of gelatinase B, provides little insight into NGAL/gelatinase B interactions as the gelatinase sequences diverge dramatically in this region.

DISCUSSION

While clearly possessing a structure homologous to other members of the lipocalin family, the NGAL calyx is distinctly larger and more open than that of other reported lipocalin structures. Distinct from other lipocalins, the inner surface of the calyx is fairly polar and carries an overall positive charge. As with many lipocalin structures, NGAL

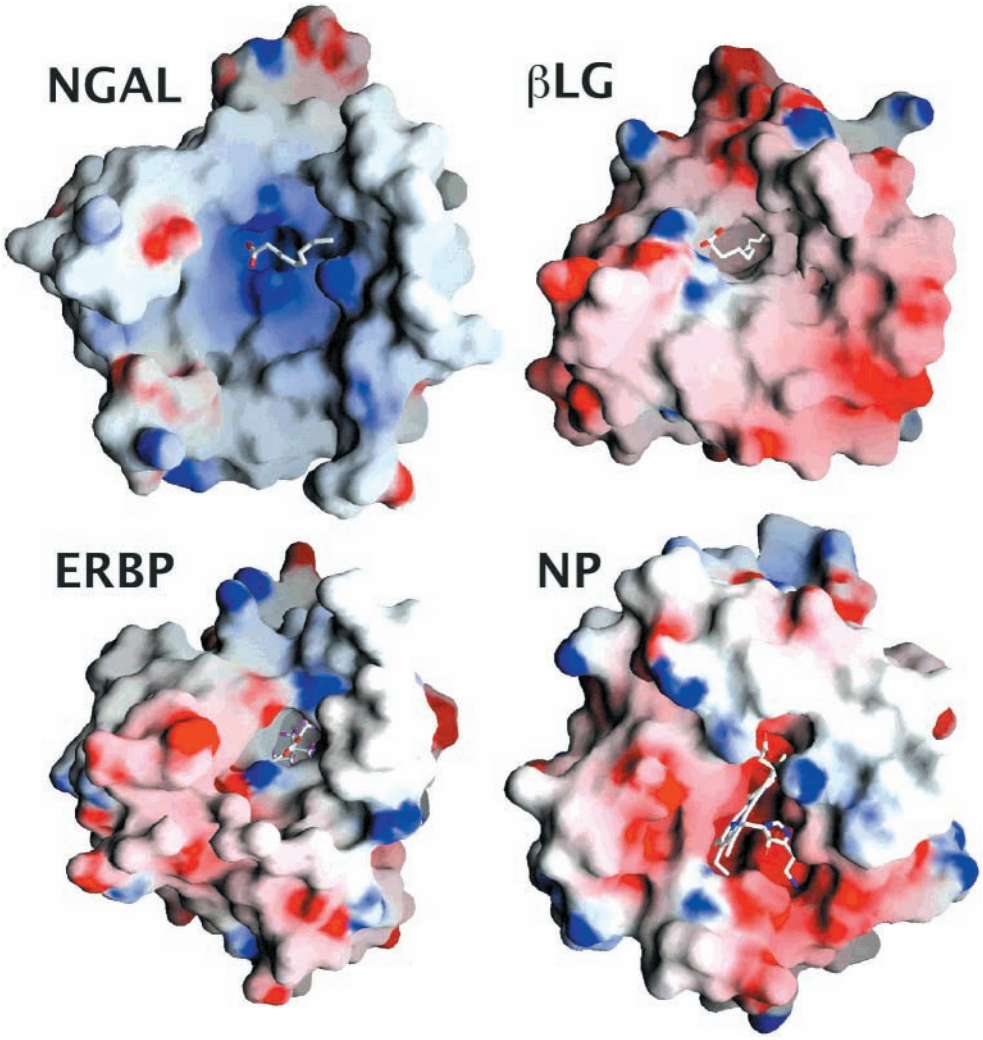

FIGURE 4: Lipocalin binding sites. GRASP (*37*) molecular surfaces, colored by electrostatic potential (blue, positive; red, negative) for NGAL, β-lactoglobulin (βLG), ERBP, and nitrophorin (NP). The bound ligands have been shown in a stick representation: NCA in NGAL, palmitate in βLG, REA in ERBP, and the heme/histamine complex in NP. The molecules have been oriented to allow views down into the calyces. The calyx of NGAL is lined with the side chains of Phe-22, Val-33, Leu-36, Ile-41, Glu-44, Gln-49, Tyr-52, Thr-54, Tyr-56, Ser-68, Leu-70, Arg-81, Phe-83, Phe-92, Leu-94, Tyr-100, Leu-103 Tyr-106, Phe-123, Lys-125, Ser-127, Tyr-132 Lys-134, Thr-136, Tyr-138, and Pro-169.

copurifies with what is likely a bound fatty acid, NCA. However, unlike the lipocalins or other fatty acid binding proteins, the NCA neither fills the NGAL calyx nor is its shape or chemical makeup particularly complementary to the pocket in which it lies. NCA is displaced by sulfate at neutral pH, with the sulfate coordinated by two lysines and an arginine, an arrangement much more suitable to binding sulfate than the aliphatic tail of a fatty acid. One of the proposed ligands, fMLF, does not bind to NGAL under either crystallization conditions and does not fit into the NCA binding site in the calyx. LTB_4, another proposed ligand, is a substituted, polyunsaturated fatty acid with hydroxyls on the 5th and 12th carbons. Placing LTB_4 into the NCA site positions the C5 hydroxyl in the sulfate binding site. While this is a better chemical fit than NCA, LTB_4 would not complement or fill the rest of the calyx. PAF, with a positively charged headgroup and long aliphatic tail, is also unlikely to be the preferred NGAL ligand. We note that the NGAL calyx is fully large enough to accommodate macromolecules, such as small proteins on the order of the size of IL-8. It is therefore likely that some macromolecule, such as a protein, is the preferred NGAL ligand.

ACKNOWLEDGMENT

We thank Barry Stoddard and Kam Zhang for helpful discussions and Thomas Earnest for assistance with data collection.

REFERENCES

1. Nielsen, B. S., Borregaard, N., Bundgaard, J. R., Timshel, S., Sehested, M., and Kjeldsen, L. (1996) *Gut 38*, 414−420.
2. Borregaard, N. (1997) *Ann. N.Y. Acad. Sci. 832*, 62−68.
3. Triebel, S., Bläser, J., Reinke, H., and Tschesche, H. (1992) *FEBS Lett. 3*, 386−388.
4. Kjeldsen, L., Johnsen, A. H., Sengeløv, H., and Borregaard, N. (1993) *J. Biol. Chem. 268*, 10425−10432.
5. Kjeldsen, L., Bainton, D. F., Sengeløv, H., and Borregaard, N. (1993) *Blood 83*, 799−807.
6. Bundgaard, J. R., Sengeløv, H., Borregaard, N., and Kjeldsen, L. (1994) *Biochem. Biophys. Res. Commun. 202*, 1468−1475.
7. Hraba-Renevey, S., Türler, H., Kress, M., Salomon, C., and Weil, R. (1989) *Oncogene 4*, 601−608.
8. Stoesz, S. P., and Gould, M. N. (1995) *Oncogene 11*, 2233−2241.
9. Flower, D. R. (1996) *Biochem. J. 318*, 1−14.
10. Pervaiz, S., and Brew, K. (1987) *FASEB J. 1*, 209−214.

11. Sengelov, H., Boulay, F., Kjeldsen, L., and Borregaard, N. (1994) *Biochem. J. 299*, 473−479.
12. Coles, M., Diercks, T., Muehlenweg, B., Bartsch, S., Zölzer, V., Tschesche, H., and Kessler, H. (1999) *J. Mol. Biol. 289*, 139−157.
13. Strong, R. K., Bratt, T., Cowland, J. B., Borregaard, N., Wiberg, F. C., and Ewald, A. J. (1998) *Acta Crystallogr. D54*, 93−95.
14. Teng, T.-Y. (1990) *J. Appl. Crystallogr. 23*, 387−391.
15. Otwinowski, Z., and Minor, W. (1996) *Methods Enzymol. 276*, 307−326.
16. McRee, D. E. (1992) *J. Mol. Graphics 10*, 44−46.
17. La Fortelle, E. d., and Bricogne, G. (1997) *Maximum-Likelihood Heavy-Atom Parameter Refinement in the MIR and MAD Methods*, Vol. 276, Academic Press, New York.
18. Collaborative Computational Project, N. (1994) *Acta Crystallogr. D50*, 760−763.
19. Kissinger, C. R., Gehlhaar, D. K., and Fogel, D. B. (1999) *Acta Crystallogr. D55*, 484−491.
20. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol. 277*, 173−208.
21. Brünger, A. T. (1987) *X-PLOR Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven and London.
22. Brünger, 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) *Acta Crystallogr. D54*, 905−921.
23. Reed, R. J. (1986) *Acta Crystallogr. A42*, 140−149.
24. Brünger, A. T. (1992) *Nature 355*, 472−475.
25. Laskowski, R. A., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) *J. Appl. Crystallogr. 26*, 283−291.
26. Colovos, C., and Yeates, T. O. (1993) *Protein Sci. 2*, 1511−1519.
27. Lüthy, R., Bowie, J. U., and Eisenberg, D. (1992) *Nature 356*, 83−85.
28. Qin, B. Y., Bewley, M. C., Creamer, L. K., Baker, H. M., Baker, E. N., and Jameson, G. B. (1998) *Biochemistry 37*, 14014−14023.
29. Newcomer, M. E. (1993) *Structure 1*, 7−18.
30. Newcomer, M. E., Pappas, R. S., and Ong, D. E. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 9223−9227.
31. Bocskei, Z., Groom, C. R., Flowers, D. R., Wright, C. E., Phillips, S. E. V., Cavaggioni, A., Findlay, J. B. C., and North, A. C. T. (1992) *Nature 360*, 186−190.
32. Brownlow, S., Morais Cabral, J. H., Cooper, R., Flower, D. R., Yewdall, S. J., Polikarpov, I., North, A. C., and Sawyer, L. (1997) *Structure 5*, 481−495.
33. Andersen, J. F., Weichsel, A., Balfour, C. A., Champagne, D. E., and Montfort, W. R. (1998) *Structure 6*, 1315−1327.
34. Benning, M. M., Smith, A. F., Wells, M. A., and Holden, H. M. (1992) *J. Mol. Biol. 228*, 208−219.
35. Morgunova, E., Tuutila, A., Bergmann, U., Isupov, M., Lindqvist, Y., Schneider, G., and Tryggvason, K. (1999) *Science 284*, 1667−1670.
36. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
37. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins: Struct., Funct., Genet. 11*, 281−296.

BI992215V