Hereditary neutropenia: dogs explain human neutrophil elastase mutations

Marshall Horwitz, Kathleen F. Benson, Zhijun Duan, Feng-Qian Li and Richard E. Person

Mutations in \textit{ELA2}, the gene encoding neutrophil elastase (NE), cause the human diseases cyclic neutropenia (CN) and severe congenital neutropenia (SCN). Numerous mutations are known, but their lack of consistent biochemical effect has proven puzzling. The recent finding that mutation of \textit{AP3B1}, which encodes the \( \beta \) subunit of adaptor protein complex 3 (AP3), is the cause of canine CN suggests a model for the molecular basis of hereditary neutropenias, involving the mistrafficking of NE: AP3 recognizes NE as a cargo protein, and their interaction implies that NE is a transmembrane protein. Computerized algorithms predict two NE transmembrane domains. Most CN mutations fall within predicted transmembrane domains and lead to excessive deposition of NE in granules, whereas SCN mutations usually disrupt the AP3 recognition sequence, resulting in excessive transport to the plasma membrane.

Neutrophils number approximately half of the circulating white blood cells, and phagocytose bacterial and fungal pathogens and coordinate the inflammatory response. Neutrophil deficiency (neutropenia) predisposes to infection. Inherited human neutropenia is rare and consists of two main forms: cyclic neutropenia (CN), also known as cyclic hematopoiesis and severe congenital neutropenia (SCN), sometimes referred to as Kostmann syndrome.

Individuals with CN display three-week oscillations in the circulating neutrophil count (Figure 1), fluctuating between near zero and almost normal levels [1,2]. Monocytes also cycle, but in the opposite phase to neutrophils. A genome-wide screen for linkage, using positional cloning, established that autosomal dominant, heterozygous mutations of the gene \textit{ELA2}, encoding the neutrophil-granule serine-protease neutrophil elastase (NE), cause all cases of CN [3].

SCN refers to non-cyclical neutropenia present from birth [4]. The bone marrow in SCN characteristically displays a pro-myelocytic arrest. Kostmann originally described a family in northern Sweden with autosomal-recessive inheritance of SCN. However, most subsequently described cases demonstrate autosomal-dominant inheritance or present sporadically (a feature of lethal, dominant disorders arising from new mutations). Furthermore, myelodysplasia (MDS) and acute myelogenous leukemia (AML) frequently complicate SCN but are absent in Kostmann’s kindred. Kostmann syndrome should, therefore, be reserved to refer to the unique features of this specific family. A candidate gene study found that most cases of SCN result from heterozygous \textit{ELA2} mutations, typically distinct from those causing CN [5] (there was a confusing report of \textit{ELA2} mutations and accelerated apoptosis in Kostmann’s original family, but this was retracted [6], pending investigation). Mutations of the gene encoding the granulocyte colony stimulating factor (G-CSF) receptor were initially reported as the cause of SCN. However, these mutations are now recognized as acquired events in the bone marrow that are common in SCN cases progressing to MDS and AML, but they are not always present, nor does their presence always signify malignancy [7].

The finding that \textit{ELA2} mutations cause CN and SCN, including all cases of SCN leading to MDS and AML (21% of SCN cases with an \textit{ELA2} mutation) [8], was a surprising result, now confirmed by multiple laboratories [8–12]. Additionally, an individual with germline \textit{ELA2} mosaicism, in whom the mutation is absent in neutrophils, has been taken as proof of causality [10] (a report suggesting that \textit{ELA2} mutations were not causative of SCN [13], was later shown to also represent germline mosaicism [14,15]).

The screening of a large number of patients with SCN lacking \textit{ELA2} mutations identified two families with heterozygous, autosomal dominant \textit{Gfi1} mutations [19]. Gfi1 is a transcriptional repressor [16] that, when deleted...
in gene-targeted mice, causes neutropenia [17,18] and results in circulating primitive myeloid cells and lymphocyte dysfunction. Gfi1 regulates the expression of many genes, including ELA2 [20], and Gfi1 mutation results in the overexpression of NE [19].

The biology of the ELA2 gene product, NE
ELA2 is transcribed only in pro-myelocytic and promonocytic progenitors in the bone marrow, but the protein, if not the transcript, persists in the cells through terminal differentiation to neutrophils and monocytes [21]. ELA2 encodes a 267-residue protein that is post-translationally processed at both termini. The N-terminus contains a 27-residue pre sequence cleaved by a signal peptidase. The protease cathepsin C, also known as dipeptidyl peptidase I (DPPI), cleaves a remaining two residue pro sequence, in a step corresponding tozymogen activation [22,23]. The C-terminus consists of a 20-residue tail that is normally excised and is required neither for granule localization nor enzymatic activity [24,25]. NE digests many substrates, including matrix components, clotting factors, immunoglobulins and complement. It is inhibited primarily by the serpin α1-antitrypsin, which forms an irreversible complex upon its cleavage by NE.

Inconsistent biochemical properties of mutant NE
Thirty different ELA2 mutations (Figure 2a) have been identified [4,8–12,26]. Mutations causing CN are generally distinct from those responsible for SCN, although the phenotypes and genotypes can overlay. Most mutations are limited to just a few families, with two exceptions: the majority of CN cases arise from intron-four splice-donor mutations, causing a deletion of ten amino acid residues (ΔV161–F170), and the most common mutations in SCN are truncations near the C-terminus.

Recombinant expression of the mutated genes in rat basophilic leukemia (RBL) cells allows for the characterization of the biochemical properties [25] of mutant forms of NE. Most, but not all, mutations causing CN or SCN reduce proteolytic activity. Mutants demonstrate no obvious change in substrate specificity and none of the mutations affects protein stability, although some of the mutant proteins might aggregate more readily. The mutant proteins differ in their susceptibility to inhibition by α1-antitrypsin, but not by more than would be expected from the loss of activity required to cleave this inhibitor. None of these properties is consistent for all mutations, regardless of whether they cause CN or SCN. Therefore, the effects of the mutations were puzzling, but the recent identification of the gene responsible for canine CN [26] has offered a surprising potential explanation.

AP3B1 mutations as the cause of canine CN
Canine CN [27] is also known as gray Collie syndrome, because it arose in Collies and the dogs have a diluted coat color. The human and canine forms of the disease differ in a number of ways: human CN lacks pigmented abnormalities, canine disease demonstrates autosomal-recessive transmission, in dogs, neutrophil counts cycle every two weeks rather than three weeks (Figure 1) and all blood cells cycle.

Using direct DNA sequencing and genetic linkage studies, canine ELA2 mutations were excluded as the cause of gray Collie syndrome. Linkage analysis of candidate genes affecting pigmentation and hematopoiesis, based on corresponding phenotypes in humans and mice, confirmed genetic linkage and linkage disequilibrium to AP3B1 [26]. The mutation responsible for gray Collie syndrome is an insertion of an adenine nucleotide residue within a tract of nine adenines in exon 20, leading to a frameshift with premature termination and absent mRNA, apparently resulting from nonsense-mediated decay [26]. AP3B1 encodes the β subunit of adaptor protein 3 (AP3). The equivalent mutation is the cause of the rare Hermansky–Pudlak syndrome type 2 (HPS2) in humans – only three families with a total of four children have been described [28–30] – as well as the pearl mouse and ruby Drosophila strains (Figure 3). Hermansky–Pudlak syndromes are heterogeneous autosomal-recessive disorders that cause cutaneous and retinal hypopigmentation and bleeding, resulting from abnormal platelet dense-granules [31,32]. Seven human and 14 mouse genes have been cloned [33], and all appear to be involved in lysosomal transport. HPS2 is the only type featuring neutropenia.

Mammalian cells contain four different heterotetrameric adaptor protein complexes, consisting of two large (α, γ, δ or ε, and β), one medium (μ) and one small (σ) subunit, which reside on the cytoplasmic surface of membrane-bound vesicles and direct subcellular trafficking of membrane cargo proteins [34]. AP3 is specifically responsible for routing cargo proteins from the trans-Golgi network to lysosomes [35]. However, there are other lysosomal transport pathways. Membrane-bound lysosomal acid phosphatase, an AP2 cargo protein, is routed indirectly to lysosomes via recycling between the endosomes and the cell surface [35]. Soluble lysosomal proteins, such as the granymes of lytic granules in lymphocytes, are targeted through the mannose 6-phosphate receptor, which specifically recognizes mannose 6-phosphate as a post-translational modification of asparagine-linked oligosaccharides [36]. AP3 cargo proteins contain a tyrosine-based motif recognized by μ subunits, or a di-leucine motif recognized by β subunits. Mutation of the β subunit in HPS2 and pearl mice is sufficient to dissociate the tetramer and induce degradation of other subunits [26,28,30,37], resulting in AP3 cargo proteins being mislocalized to the plasma membrane [28,30,38]. Granymes are correctly targeted through the intact mannose-6-phosphate receptor pathway, but the lytic granules of lymphocytes from patients with HPS2 are immotile and unable to dock at the cell surface for secretion [30]. A multitude of adaptor protein mutations have been cataloged in yeast, flies and mice [31].

Interaction between the gene products responsible for human and canine CN
Given the similarities between canine and human CN, an obvious question is whether NE interacts with AP3? Specifically, does NE bind to β3a or β3a, the subunits responsible for recognizing cargo proteins? A yeast two-hybrid assay established for testing adaptor protein subunit and cargo protein interactions [39],
indicates that µ3a interacts with NE via a tyrosine-based recognition signal upstream of the C-terminal tail, but only after the C-terminal tail is removed [26]. Because the C-terminus blocks interactions with µ3a, normal C-terminal processing should permit recognition of NE by AP3.

The evidence, therefore, indicates that NE is an AP3 cargo protein [26]: mutations in either NE or the β subunit of AP3 result in a similar disease phenotype; NE physically associates with AP3 in vitro; the localization of NE within lysosome-like granules is similar to the distribution of other AP3 cargo proteins; and gray Collies have intact ELA2 genes, but nearly undetectable levels of NE protein [26]. Yet, there is a surprising implication of this interaction. Because cargo proteins are on the inside of vesicles, and adaptor proteins are on the outside, cargo
proteins, including integrins, clotting factor V and some other proteins that the apices of these loops immerse into the membrane, as is the case for clotting factor V and some other proteins that associate with membranes.

**NE mutations aligning with predicted transmembrane regions**

Previously, NE was recognized as a soluble protein, and its crystal structure supports its behavior as a textbook serine protease [40]. However, would computerized algorithms identifying transmembrane domains detect their presence in human NE? Surprisingly, most programs predict two transmembrane domains [26], each of which is bracketed by disulfide bonds (Figure 2a). Strikingly, when the mutations responsible for hereditary neutropenia are superimposed on the predicted transmembrane domains, a pattern emerges: mutations causing either CN or SCN generally align with predicted transmembrane segments or disrupt the disulfide bonds bracketing them, and the cluster of chain-terminating mutations causing SCN deletes the μ3a recognition signal (along with the C-terminus). Some programs predict a third transmembrane segment (TM-cryptic) whose formation is facilitated by an SCN mutation (S97L) in a region near the location of two other mutations (P110L and 93 + PQ). The inference is that NE is a transmembrane protein. The somewhat controversial notion that a transmembrane form of NE is in equilibrium with a soluble form in which the transmembrane segments are constrained into disulfide-bonded loops (Figure 2b) has been proposed [26]. It is also possible that the apices of these loops immerse into the membrane, as is the case for clotting factor V and some other proteins that associate with membranes.

This should not be surprising, as considerable evidence places NE on the plasma membrane and in other intracellular membranes, as well as within the lumen of granules. The first purification of NE was from membranes [41] and soluble NE can be extracted from purified plasma membranes [42]. Electron microscopy utilizing antibodies against NE [43] and labeled substrates [44] detects NE in the plasma membrane, Golgi and nuclear envelope. Membrane forms of NE cleave membrane proteins, including integrins [45], clotting factor V [46] and the transferrin receptor [47]. NE might serve as a membrane co-receptor for HIV infection of macrophages [48]. Selective cell permeabilization studies support the transmembrane topology of NE [26]. An explanation for how a membrane form of NE equilibrates into a soluble form within the lumen is lacking. Nevertheless, a mixture of purified NE and purified plasma membranes demonstrates reversible partitioning between soluble and membrane bound forms [42], and there is evidence that the membrane association takes place through genuine hydrophobic, rather than electrostatic, interaction [49].

---

**Box 1. A disease model based on mistrafficking of neutrophil elastase (NE)**

Ordinarily, NE containing an intact C-terminal tail exists, at least transiently, in a transmembrane configuration in the trans-Golgi network, with its termini extruding into the cytoplasm (Figure 1a). The carboxyl tail blocks interactions with μ3a, but this is usually cleaved, permitting NE interaction with μ3a. Adaptor protein 3 (AP3) then shuttles the tail-less form of NE to granules, where it subsequently equilibrates into a soluble form within the granule lumen. Less frequently, some NE molecules retain the C-terminus, remain membrane bound and traffic to other membrane compartments (such as the plasma membrane) which is the default destination for cargo proteins in the absence of AP3 [28,30,38]. (Figure 1 illustrates a model in which AP3 attaches to a preformed membrane bud through interaction with the cytoplasmic tail of the cargo protein [50]. Other data suggest that AP3 attaches to a membrane independently of cargo protein recognition and is directly involved in the budding step [51–53]. In either case, complex formation with AP3 might favor an equilibrium shift in NE from the soluble to the transmembrane form.)

In this model, the effect of the mutations would be as follows:

(i) Severe congenital neutropenia (SCN) mutations disrupting the μ3a recognition signal send NE to the default destination, the plasma membrane (Figure 1b).

(ii) In Hermansky–Pudlak syndrome type 2 (HPS2), μ3a mutations disrupting the AP3 tetramer also direct NE to the plasma membrane, thus causing non-cyclic neutropenia (Figure 1c). (μ3a mutations also cause the gray Collie syndrome, but the neutropenia is inextricably cyclic in dogs.)

(iii) Mutations in the transcriptional repressor Gfi1 lead to overexpression of NE (Figure 1d); normal AP3-based granular transport pathways become overwhelmed, causing excessive NE, with an intact C-terminus, to locate to the plasma membrane, thus causing SCN. Some patients with SCN lacking mutations in the ELA2 coding sequence demonstrate DNA sequence changes in the promoter that upregulate ELA2 expression and that might act in a similar manner to Gfi1 mutations [54], although genetic evidence supporting causality is inconclusive.

(iv) Most mutations capable of causing CN perturb transmembrane segments, thereby favoring a shift towards a soluble form of NE, localizing within granules (Figure 1e). Mutations deleting the C-terminal tail but leaving the μ3a recognition signal intact are not known, but the model predicts that they would facilitate AP3 interaction and granule localization, thus causing CN.
Figure I. Cellular model for hereditary neutropenia. Gray arrows show normal pathways and blue arrows indicate pathological trafficking. Abbreviations: AP3, adaptor protein 3; CN, cyclic neutropenia; HPS2, Hermansky–Pudlak syndrome 2; NE, neutrophil elastase; SCN, severe congenital neutropenia; TGN, trans-Golgi network.
To determine whether NE mutations result in membrane mis trafficking, the distribution of mutants expressed in RBL cells has been tracked, using both immunofluorescent staining and density gradient cell fractionation [26]. Wild-type enzyme predominantly localizes to granules and appears as the mature isoform from which the C-terminus has been removed. A smaller proportion of NE contains the intact C-terminus, and localizes to the plasma membrane. The distribution of the mutations generally provides a correlation between genotype and phenotype. Mutations disrupting either of the predicted transmembrane domains typically accumulate as the isoform lacking the C-terminus, localize within granules and usually cause CN. By contrast, mutations that disrupt the μ3a recognition signal preferentially localize in the plasma membrane and seem to exclusively cause SCN. A disease model for CN and SCN, based on the trafficking of NE, has been proposed (Box 1) [26]. Thus, we hypothesize that the two forms of hereditary neutropenia, CN and SCN, represent an imbalance between alternative intracellular transport pathways for NE (Box 2).

**Lack of correspondence among species**

It is not readily apparent why human disease differs from animal models. Knock-in gene targeting to introduce a human SCN mutation (V72M) into the corresponding position of normal mouse ELA2 failed to produce an abnormal phenotype [55]. (Ironically, V72M is the only mutation unaccounted for by the categorization scheme of Figure 2a and, retrospectively, might have been an unlucky choice.) The absence of μ3a causes cyclic pancytopenia in gray Collie dogs, non-cyclic neutropenia in humans with HPS2 and no cytopenias in pearl mice [37] (although cyclic neutropenia could potentially go undetected in a mouse). One possible explanation is species-specific differences in the NE sequence; it is worth noting that one mutation, V190I, changes a residue unique in humans to a residue conserved between mouse and dog.

**Extension to other human neutropenic syndromes**

An appeal of this model is its potential extension to include other neutropenic syndromes. Barth syndrome is a sex-linked recessive disorder of cyclic neutropenia and dilated cardiomyopathy, caused [56] by mutations to G4.5/TAZ. This gene encodes an acyl transferase involved in the synthesis of cardiolipin [57], a mitochondrial membrane lipid. Cohen syndrome comprises autosomal recessive inheritance of mental retardation, dysmorphic features and neutropenia [58]. The responsible gene, COH1, encodes a protein with homology to a yeast protein, VPS13, involved in vesicle sorting and intracellular protein transport [59]. Defects in membrane composition and vesicle transport raise the intriguing possibility that mislocalization of NE is responsible for neutropenia in these two syndromes.

**Pathogenic targets of NE**

What are the substrates of mislocalized NE? Any answer to this question must ultimately account for the periodicity observed in CN. Many groups have proposed that oscillations suggest a feedback circuit in which mature neutrophils inhibit progenitor cells [60]. Inhibition of progenitors leads to the loss of successive cohorts of maturing cells, eventually depleting the generation of cells producing the inhibitory signal, thus allowing the pattern to repeat. It is, therefore, expected that the targets of the mutant forms of NE ultimately feedback to regulate ELA2 expression. There are several candidate NE substrates at the cell surface and in the extracellular environment, including G-CSF [61–63], the G-CSF receptor [62] and the c-KIT receptor [64].

NE also interacts with and cleaves a novel Notch family member, N2N, as well as Notch2 [65]. Notch proteins are cell-surface receptors specifying binary cell fate in a cell-autonomous manner, and whose proteolytic cleavage yields an intracellular fragment with transcriptional activity [66]. Involvement of Notch signaling is an attractive prospect for three reasons. First, a feature of hereditary neutropenia is the reciprocal relationship between the numbers of monocytes and neutrophils, corresponding to the alternative fates of myeloid progenitor cells. Second, as a transcription factor, Notch could conceivably feedback to govern ELA2 expression. Third, in germline mosaicism in SCN, the ELA2 mutation is absent in circulating neutrophils, indicating that mutations act in a cell-autonomous fashion [10].

It appears that NE is normally present within the nucleus [67], and that neutrophenic mutations result in accumulation within the nuclear envelope [26,68]. A recent finding is that the PML/RARA fusion gene, the product of the chromosome t(15;17) translocation in promyelocytic leukemia, is cleaved by NE in the nucleus and that NE-deficient mice expressing a PML/RARA transgene are resistant to leukemia [69].

There is also the possibility that NE in disease might have another activity altogether: the activities of membrane-associated NE might be altered [65,70]; predicted TM domains include portions of the catalytic triad of the enzyme (formed by the amino acids His, Asp, and Ser); and residues outside of the triad might participate in catalysis [71].

**Concluding remarks**

NE has a significant role in myelopoiesis, and its previously underappreciated localization in cellular compartments other than granules, with attendant implications for NE as a membrane protein, opens a search for identifying important regulatory targets. This has
far-reaching implications for serine protease biochemistry, subcellular trafficking, hematopoiesis and the pathogenesis of neutropenia and leukemia. Outstanding questions remain (Box 3) and future studies will aim to answer them.

Acknowledgements
We thank David Ehlerl for assistance preparing figures, the estate of Dr Lee Ford for access to her photo archive, and Dr Richard Swank for mouse and Ginny Cuneo for dog photographs. Supported by grants (to M.H.) from the NIH (DK55820, DK58161) and Burroughs-Wellcome Fund (SART-1002189).

References
9. Ancill, P.J. et al. (2001) Mutations in the ELA2 gene encoding neutrophil elastase are present in most patients with sporadic severe congenital neutropenia but only in some patients with the familial form of the disease. Blood 98, 2645–2650
33. Li, W. et al. (2003) Hermansky–Pudlak syndrome type 7 (HPS-7) results from mutant dysbindin, a member of the biogenesis of lysosome-related organelles complex 1 (BLOC-1). Nat. Genet. 35, 84–89
53 Faundez, V. et al. (1998) A function for the AP3 coat complex in synaptic vesicle formation from endosomes. *Cell* 93, 423–432
55 Grenda, D.S. et al. (2002) Mice expressing a neutrophil elastase mutation derived from patients with severe congenital neutropenia have normal granulopoiesis. *Blood* 100, 3221–3228