The Met receptor tyrosine kinase and its ligand hepatocyte growth factor (HGF) play an important role in mediating both tumor progression and tissue regeneration. The N-terminal and first Kringle domains (NK1) of HGF comprise a naturally occurring splice variant that retains the ability to activate the Met receptor. However, NK1 is a weak agonist and is relatively unstable, limiting its therapeutic potential. Here, we engineered NK1 mutants with improved biochemical and biophysical properties that function as Met receptor agonists or antagonists. We first engineered NK1 for increased stability and recombinant expression yield using directed evolution. The NK1 variants isolated from our library screens acted as weak Met receptor antagonists due to a mutation at the NK1 homodimerization interface. We introduced point mutations that restored this NK1 homodimerization interface to create an agonistic ligand, or that further disrupted this interface to create more effective antagonists. The rationally engineered antagonists exhibited melting temperatures up to approximately 64 °C, a 15 °C improvement over antagonists derived from wild-type NK1, and approximately 40-fold improvement in expression yield. Next, we created disulfide-linked NK1 homodimers through introduction of an N-terminal cysteine residue. These covalent dimers exhibited nearly an order of magnitude improved agonistic activity compared to wild-type NK1, approaching the activity of full-length HGF. Moreover, covalent NK1 dimers formed from agonistic or antagonistic monomeric subunits elicited similar activity, further signifying that NK1 dimerization mediates agonistic activity. These engineered NK1 proteins are promising candidates for therapeutic development and will be useful tools for further exploring determinants of Met receptor activation.

Results

NK1 Engineering. HGF exists as two different isoforms, isoform 1 (I1) and isoform 3 (I3), that are identical in sequence, except for a five amino acid deletion in the first Kringle domain (K1) of I3 compared to I1. We expressed NK1 fragments corresponding to I1 or I3 on the yeast cell surface as genetic fusions to the native protein ligands are promising starting points for developing receptor agonists or antagonists (6). HGF is an approximately 80 kDa multidomain protein composed of an N-terminal hairpin domain, four Kringle domains, and a C-terminal serine protease homology domain (Fig. S1) (7). Although HGF is thought to bind to the Met receptor Sema domain (8–10), a critical role for the immunoglobulin-like (IPT) domains of Met has also been reported (11). HGF is secreted as an inactive single-chain protein that must be activated through proteolytic cleavage into a disulfide-linked α/β heterodimer. Point mutations that disrupt the proteolytic processing of HGF (12) or the conformational maturation following cleavage (13) transform full-length HGF from a Met agonist into an antagonist. By contrast, a synthetic HGF fragment comprising the N-terminal domain and all four Kringle domains (NK4) functions as a Met antagonist without further modification (14).

The NK1 fragment of HGF (comprising the N-terminal and first Kringle domains) is a naturally occurring HGF splice variant that retains binding to Met, but elicits much weaker Met receptor activation compared to full-length HGF. NK1 binding to Met is enhanced by heparan-sulfate proteoglycans (HSPGs) (15). Met receptor dimerization and activation is mediated through a heparin-induced NK1 homodimer (16, 17), and these molecular events appear to underlie the mechanism of HGF-Met activation (18, 19). The NK1 fragment can be transformed from a weak agonist into an antagonist through point mutations targeted to its homodimerization interface (19). NK2, another naturally occurring HGF splice variant, comprises the N-terminal and first two Kringle domains of HGF. NK2 can be transformed from an antagonist to an agonist by disrupting the NK1-K2 intraprotomer interaction and then can be turned back into an antagonist by introducing the same point mutations that block NK1 homodimerization (20).

The HGF variants described above have provided important insights into HGF-Met structure and function; however, they generally suffer from low recombinant expression yield or stability, limiting their practical use and study. For example, although NK1 is one of the smallest HGF fragments, we found that it possesses relatively low stability, and previous reports have noted difficulty with expression (19, 21). Here, we used directed evolution to engineer NK1 mutants with substantial improvements in stability and recombinant expression yield compared to wild-type NK1. Next, we used the mechanistic basis of NK1-Met activation to rationally introduce point mutations into these NK1 mutants to generate agonistic or antagonistic variants. In addition, we created covalent NK1 homodimers that exhibited remarkable improvements in agonistic activity compared to wild-type NK1, with activity levels approaching that of full-length HGF.
agglutinin mating protein Aga2p (Fig. S1 B and C). NK1 I1 was expressed on the yeast cell surface, but exhibited only weak levels of binding to soluble, recombinant Met extracellular domain (Met-Fc) (Fig. S1D). Similar results were obtained in the presence of heparin (Fig. S1D), which has been shown to enhance the affinity of the NK1-Met interaction (15). Binding levels were indistinguishable for 20 or 200 nM Alexa-488 labeled Met-Fc (Met-Fc A488) and remained unchanged after heating to 70 °C (Fig. S1E), suggesting that the low levels of binding observed was nonspecific. Similar behavior was observed for NK1 I3.

We performed two rounds of directed evolution to identify NK1 mutants with improvements in stability and Met binding affinity as outlined in Fig. 1. Using error-prone PCR, we created a first-generation mutant library (M1) of approximately $3 \times 10^7$ yeast transformants. The library contained a relatively equal mixture of mutants based on NK1 I1 or NK1 I3, as confirmed by DNA sequencing. To facilitate folding of NK1 mutants, we conducted two rounds of library screening after inducing for yeast cell surface expression at 20 °C. High-throughput FACS was used to isolate functional NK1 mutants that bound to Met-Fc A488 (Fig. 1B). Subsequent library sorts were conducted in parallel, using either 20 or 30 °C expression temperatures (Fig. 1 C and D), with the goal of screening for NK1 mutants with improved stability using the higher temperature. After six or seven rounds of library sorting (seven rounds with 20 °C expression temperature, or two rounds with 20 °C followed by four rounds with 30 °C expression temperature) the library was enriched with mutants that bound to soluble Met-Fc A488.

The pools of mutants from the final sort rounds above were randomly mutated by error-prone PCR to generate a second-generation mutant library (M2) of approximately $8 \times 10^7$ yeast transformants. The first two rounds of sorting of this library were conducted using a 20 °C expression temperature to recover mutants that bound to soluble Met-Fc A488 (Fig. 1E). For subsequent sort rounds, we screened in parallel for NK1 mutants with increased Met binding affinity using decreasing concentrations of soluble Met-Fc A488 (Fig. 1 F and G), or for NK1 mutants with increased stability using expression at elevated temperatures (37 °C) (Fig. 1 H and I). Ultimately, we obtained a pool of NK1 mutants that had been screened for improvements in Met binding affinity, and a second pool that had been screened for improved stability while retaining Met binding capability.

**Sequences of NK1 Mutants.** Following the second round of directed evolution, we sequenced eight randomly selected clones from each of the final two sort rounds from the affinity- and stability-sorted pools. All 32 clones analyzed contained the NK1 I1 sequence and none were based on NK1 I3; thus, in subsequent studies we used wild-type and mutant NK1 proteins based on NK1 I1. Interestingly, there was substantial overlap in mutations identified from the affinity- and stability-sorted library pools, with eight mutations present in over half of the individual NK1 mutants. These eight dominant mutations are distributed across the NK1 surface as shown in Fig. 2. None of the individual mutants contained all eight of the most common mutations; however, one variant from the affinity-sorted pool, referred to as M2.1, contained five of the eight most frequent mutations and no additional mutations (K62E, K127D, K137R, K170E, and N193D). We added the remaining three mutations (Q95R, K132N, and Q173R) to this variant to generate the NK1 mutant we term M2.2 (Table 1).

**Binding of NK1 Proteins to Cellular Met Receptor.** NK1 binds both to Met and heparin/HSPGs, which are widely coexpressed on the mammalian cell surface. NK1 mutations located in the proximity of the heparin binding sites (K62E, Q95R, K132N, and K170E) still allowed for strong, albeit decreased, heparin binding compared to wild-type NK1 (Fig. S2A). Therefore, we used two independent approaches to measure NK1 binding to cellular Met receptor, taking care to avoid the complicating factors of NK1 binding to HSPGs. In one approach, we used BaF3 cells, a murine pre-B cell line that is devoid of HSPG expression (15), that have been stably transfected with Met (termed BaF3-Met cells) (22). Wild-type NK1 and the mutant M2.2 bound to BaF3-Met cells, but not to untransfected BaF3 cells (Fig. S3A). Binding was inhibited by preincubation with the neutralizing Met antibody AF276 (R&D Systems), demonstrating Met receptor binding specificity (Fig. S3B). In a second approach, we blocked HSPGs on the surface of A549 human lung carcinoma cells by preincubation with basic fibroblast growth factor (bFGF; also known as FGF-2), which binds to HSPGs. Following preincubation with bFGF, no decrease in binding was observed with AF276, whereas M2.2 binding was diminished, demonstrating binding of M2.2 to HSPGs on the mammalian cell surface (Fig. S3C). M2.2 binding...
Wild-type and mutant NK1 proteins each unfolded irreversibly to discriminate folded and unfolded NK1 proteins using CD.

Using these assays, M2.2 exhibited similar binding affinities to both BaF3-Met cells and A549 cells, with equilibrium binding constant ($K_D$) values of 22 ± 5 nM and 30 ± 10 nM, respectively (Fig. S4 A and B). Wild-type NK1 bound to Met expressed on BaF3-Met cells with a $K_D$ of 16 ± 4 nM (Fig. S4A), however, bound nonspecifically to A549 cells even in the presence of bFGF. M2.1 exhibited minimal binding to Met in both assays, which was unexpected given that M2.1 was isolated from the affinity-sorted pool of NK1 mutants in the second round of directed evolution. In contrast, both wild-type NK1 and M2.2 bound to Met with affinities in the low nanomolar range.

**NK1 Mutants Exhibit Enhanced Stability over Wild-Type NK1.** To retain its structural integrity, we found that wild-type NK1 must be maintained in buffer containing high salt concentrations (>200–300 mM NaCl). As further evidence of this requirement, wild-type NK1 exhibited a broad, delayed elution profile on size-exclusion chromatography with buffer containing moderate salt concentration (137 mM NaCl) (Fig. 3A, Inset), suggesting unfolding and/or nonspecific binding to the column under these conditions. In contrast, M2.1 and M2.2 eluted as a single, sharp peak on size-exclusion chromatography under similar moderate salt conditions (Fig. 3A).

We tested the thermal stability of the NK1 proteins using two independent approaches. In the first approach, binding measurements of yeast-displayed M2.1 or M2.2 to Met-Fc A488 after heating to various temperatures resulted in $T_m$ values of 61 ± 1°C and 61.4 ± 0.7°C, respectively (Fig. S5A). It was not possible to monitor stability of yeast-displayed wild-type NK1 in this manner because it was not functionally expressed on the yeast cell surface (Fig. S1 D and E). In the second approach, we used CD scans to monitor secondary structure unfolding as a function of temperature. CD spectra of soluble, recombinant M2.1 and M2.2 were indistinguishable from wild-type NK1 (Fig. S5B) and were in good agreement with previous reports for wild-type NK1 (23). A spectra of wild-type NK1 at 80°C resembled that of denatured protein (Fig. S5B) (24), demonstrating the ability to discriminate folded and unfolded NK1 proteins using CD. Wild-type and mutant NK1 proteins each unfolded irreversibly in variable temperature CD scans, and unfolding at 208 nm demonstrated higher thermal stability for both M2.1 and M2.2 ($T_m = 63.9 ± 0.5°C$ and $69 ± 1°C$, respectively) compared to wild-type NK1 ($T_m = 50.7 ± 0.2°C$) (Fig. 3B). These results were confirmed by monitoring the unfolding of M2.1 near the local maximum at 236 nm, which was in agreement with the $T_m$ value measured at 208 nm (approximately 65°C and 64°C, respectively).

**Point Mutations at the NK1 Homodimerization Interface.** Residue N127 lies within the linker region connecting the N and K1 domains (Fig. 2). The side chain of this asparagine residue forms two hydrogen bonds central to the NK1 homodimer, and these stabilizing interactions are important to the agonistic activity of NK1 (17, 19). In our library-isolated variants, we observed frequent mutation of this asparagine to aspartate (N127D) (Table 1). To probe the effects of the N127D mutation on biological activity, we generated a series of point mutants at this position. An alanine residue was previously shown to transform wild-type NK1 from an agonist into an antagonist by disrupting stabilizing interactions of the NK1 homodimer (19). We also explored the effects of mutation to lysine and arginine at this position, which introduce steric and electrostatic obstructions through bulky, charged side chains. In addition, we generated the point mutant D127N, which reverts this position back to the wild-type asparagine residue. Within the context of M2.2, which contains the N127D mutation, we refer to these mutations as D127A, D127K, and D127R. Importantly, each of these mutants retained the high thermal stability associated with M2.2 (Table S1). In general, we observed a correlation between thermal stability, as measured by CD, and yeast surface expression levels (Fig. 3C) or percent yeast surface expression (Fig. 3D). A similar relationship was observed for yeast surface expression properties and recombinant expression yield (Fig. S5 C and D).

**Characterization of NK1 Mutants as Met Receptor Agonists or Antagonists.** The mutants M2.2 D127A, D127K, and D127R did not induce Met activation, as measured by scatter (Fig. 4A and Fig. S6A) or uPA activation (Fig. S6B) in Madine–Darby canine kidney (MDCK) cells. The unmodified M2.2 variant, which

Table 1. Sequences of NK1 mutants M2.1 and M2.2 compared to human HGF (huHGF)

<table>
<thead>
<tr>
<th></th>
<th>N domain</th>
<th>Linker*</th>
<th>K1 domain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>65 95</td>
<td>127</td>
<td>132 137 170 173 193</td>
</tr>
<tr>
<td>huHGF</td>
<td>K Q</td>
<td>N K K Q N</td>
<td></td>
</tr>
<tr>
<td>M2.1</td>
<td>E D</td>
<td>R E D</td>
<td>N R E R D</td>
</tr>
<tr>
<td>M2.2</td>
<td>E R D</td>
<td>N R E R D</td>
<td></td>
</tr>
</tbody>
</table>

Blank spaces indicate same residue as huHGF. HGF isoform 1 numbering scheme is used.

*Linker region connecting N and K1 domains of NK1.
contains the N127D mutation, exhibited weak (Fig. S6A) or no agonistic activity (Fig. 4A and Fig. S6B). In contrast, reversion of position 127 to the wild-type asparagine residue (M2.2 D127N) resulted in agonistic activity in both MDCK scatter (Fig. 4A and Fig. S6A) and uPA assays (Fig. S6B). The activity of M2.2 D127N was similar to that of wild-type NK1, and both showed enhanced activity in the presence of soluble heparin (Fig. S6C). In comparison, M2.2 D127A, D127K, and D127R did not exhibit agonistic activity in these assays either in the presence or absence of heparin (Fig. 4A and Fig. S6A–C).

We next tested the ability of these mutants to inhibit HGF-induced Met activation. M2.2 D127N did not inhibit HGF-induced activity, providing further evidence of its function as a Met receptor agonist (Fig. 4B and Fig. S7). M2.2 mutants D127A, D127K, and D127R exhibited weak or minimal inhibition of HGF-induced MDCK cell scatter in the absence of soluble heparin (Fig. S7, Top). In contrast, strong antagonistic activity was observed with the addition of 2 μM heparin (Fig. S7, Bottom), likely due to a role of heparin in enhancing the NK1-Met binding interaction. Preformulating the NK1 mutants with a 2:1 molar ratio of heparin:NK1 was sufficient to confer this antagonistic activity (Fig. 4B). Unmodified M2.2 (M2.2 N127D) exhibited only weak antagonistic activity with a 2:1 molar ratio of heparin (Fig. 4B), supporting the utility of the rationally engineered point mutations. The antagonistic activity of M2.2 D127K is similar to that of the previously reported antagonist NK1 N127A (Fig. 4B). However, the M2.2 D127 A/K/R mutants possess markedly improved stability and expression compared to NK1 N127A, namely lower salt-dependent stability, an increased Tm of approximately 15 °C, and approximately 40-fold improved recombinant expression yield, which are all attractive properties for therapeutic development.

**Covalent M2.2 D127N and M2.2 D127K Homodimers Exhibit Potent Agonistic Activity.** We generated NK1 homodimers by coupling monomers through a direct disulfide bond. We hypothesized that such covalent homodimers would elicit enhanced agonistic activity compared to wild-type NK1, which relies on noncovalent heparin-induced dimerization for receptor activation (15, 19). Based on the proximity of the N termini of the two protomers in the NK1 homodimer (Fig. S8A), a free cysteine residue was introduced at the N terminus of M2.2 D127N. Interestingly, spontaneous formation of disulfide-linked dimers was observed upon recombinant expression in yeast (Fig. S8B). These cysteine-linked M2.2 D127N dimers (termed cdD127N) were purified from monomeric species by size-exclusion chromatography and used directly in biological assays. We also generated a cysteine-linked M2.2 D127K dimer (termed cdD127K) to test the biological activity of covalent homodimers based on an antagonistic monomeric subunit. The cdD127N and cdD127K proteins migrated as monomers when analyzed by reducing SDS-PAGE (Fig. S8C), confirming that dimers are formed through a direct disulfide bond. Remarkably, both cdD127N and cdD127K induced MDCK cell scatter at an order of magnitude lower concentration than the M2.2 D127N monomer (Fig. 5A and Fig. S8D), which shows comparable agonistic activity to wild-type NK1 (Fig. S6C). The cysteine-linked dimers also elicited similar increases in uPA activation compared to the M2.2 D127N monomer (Fig. 5B and C).

**Discussion**

Here, we used a combination of approaches to develop therapeutic protein ligands with improved biochemical and biophysical properties. Directed evolution of NK1 identified mutations that...
conferred improved stability and recombinant expression yield while maintaining receptor binding affinity, and subsequent rational engineering studies generated Met receptor agonists or antagonists.

**Receptor Binding Affinity.** We developed two strategies to measure binding of NK1 to cellular Met receptor, taking care to avoid complicating factors of contributions by HS PGs. We found the binding affinity of wild-type NK1 to cellular Met receptor to be in the low nanomolar range (16 ± 4 nM). Heparin has been reported to enhance NK1-Met binding affinity (15, 19, 20), but here we report reasonably strong NK1-Met binding in the absence of heparin. This affinity is considerably stronger than previous measurements that used recombinant forms of the Met extracellular domain (20, 25), suggesting that Met may adopt different conformations when expressed on the mammalian cell surface. It is also possible that the instability of wild-type NK1 has made accurate determination of its Met binding affinity challenging. Interestingly, wild-type NK1 reached a higher maximum binding plateau than M2.2 (Fig. S4A). This could be due to differences in kinetic dissociation rate constants, perhaps with M2.2 unbinding more quickly during the secondary antibody incubation and wash steps. Alternatively, an intriguing possibility is that Met receptor binding could induce homodimerization of NK1, but not M2.2, resulting in an increased binding signal for wild-type NK1 due to higher binding stoichiometry. This also would suggest that both Met binding and heparin binding are involved in promoting NK1 homodimerization.

Surprisingly, the NK1 mutants we identified did not exhibit increased affinity over wild-type NK1. The low starting stability of wild-type NK1 likely resulted in a requirement for substantial stability evolution, and thus subsequent screens for improved affinity may have served only to restore the native affinity of wild-type NK1. For example, M2.1 and M2.2 exhibited somewhat similar stabilities (Tm = 64 °C and 69 °C, respectively), but M2.2 had substantially higher Met binding affinity than M2.1 (Fig. S4). Nevertheless, these mutants provide a good starting point for further affinity maturation studies, and the approaches outlined here can be used to monitor Met receptor interactions in their native cellular context.

**Stability Engineering and Relationship to Soluble Expression Yield.** Though wild-type NK1 exhibited a Tm of approximately 51 °C, we had difficulties with its expression, purification, and storage, and substantial troubleshooting and optimization were required in order to obtain functional protein. As an example, functional production and purification of wild-type NK1 required buffers to be autoclaved to prevent protein degradation. In addition, wild-type NK1 required high salt concentrations (>200–300 mM NaCl) to prevent destabilization and denaturation. By comparison, NK1 mutants M2.1 and M2.2 were considerably less sensitive to buffer conditions and retained their activity in buffers with lower salt concentrations.

The enhanced stability of M2.1 was surprising given that it was isolated from the affinity-sorted pool of mutants. Because wild-type NK1 is not functionally expressed on the yeast cell surface, it is possible that simply evolving for functional expression resulted in enhanced protein stability. Comparison of wild-type NK1 and each of the engineered NK1 mutants revealed a correlation between yeast surface expression levels and thermal stability (Fig. 3C). The proteins were partitioned into populations of low stability/low yeast surface expression (comprising wild-type NK1 and NK1 N127A) or high stability/high yeast surface expression (comprising the stability-enhanced NK1 mutants) (Fig. 3C). In yeast cell surface display, an intrinsic negative yeast population is always present that does not express protein (Fig. 1 and Fig. S1D); this is generally attributed to plasmid loss (26), but a role for protein stability has also been speculated (27). In support of the latter, we observed a similar correlation between thermal stability and the percentage of yeast that express protein on their surface, with the strongest trends observed for expression at 30 and 37 °C (Fig. 3D). Additionally, similar trends were observed between recombinant yield and yeast surface expression levels (Fig. S5C) or percent yeast surface expression (Fig. S5D). Taken together, these results demonstrate the power of yeast induction temperature to impart library screening stringency on thermal stability and recombinant yield in directed evolution studies. Consequently, the engineered NK1 mutant M2.2 was produced at substantially higher yield than wild-type NK1 (30 mg/L for M2.2, compared to 3.5 mg/L for NK1). These results are in general agreement with previous studies that report a strong correlation between yeast surface expression levels and soluble secretion efficiency (28, 29) or between yeast surface expression level and thermal stability (30) for proteins with immunoglobulin folds.

**Receptor Activation and Mechanistic Insights.** Interestingly, 7 out of the 11 commonly occurring NK1 mutations were changed to the corresponding residues in the HGF family member macrophage stimulating protein (MSP), HGF orthologs, or other Kringle domains (Table S2). Such mutations are often found in improved variants isolated from combinatorial libraries (31) and implies that for a given tertiary structure nature selects particular residues for protein folding, expression, and stability. Residues K62, K132, and K170 are located at the heparin binding sites within NK1. The mutation to the opposite charge (K62E, K170E) or a neutrally charged residue (K132N) within mutant M2.2 may improve stability by allowing the energetic penalty associated with clustering positive charges to form the heparin binding site (Fig. S2B–E). The Q173R mutation occurs at a position previously reported to be important for interaction with Met (32), and K137R lies at the N1-K1 intradomain interface and may potentially form a cross-domain interaction that stabilizes NK1 conformational flexibility.

Restoration of the native NK1 homodimerization interface in the stability-enhanced mutant M2.2 D127N results in essentially identical agonistic activity to wild-type NK1, and disruption of this interface by the M2.2 D127A/K/and R mutants results in efficient receptor antagonism, further supporting the physiologic relevance of NK1-NK1 crystallographic interface (17, 19, 20). The improved NK1 mutants described here could support further development of other Met agonists that are comprised partially or wholly from the unstable wild-type NK1 fragment (33–35), and given their enhanced stability have potential to be readily incorporated into biomaterials for therapeutic applications (36). In addition, these engineered NK1 agonists and antagonists will be important tools for further elucidating determinants of Met binding and activation. As previous studies point to the biological relevance of the NK1-NK1 dimerization interface (8, 18, 19), information into NK1 structure and function will provide direct insights into the mechanism of full-length HGF-Met activation. Interestingly, even though the individual M2.2 D127K monomeric subunits have antagonistic activity, cddD127K is a potent agonist with similar activity to cddD127N, highlighting the importance of a NK1 homodimer in mediating activation. Moreover, the ability to transform the antagonistic M2.2 D127K monomer into an agonist by covalent cross-linking implies a receptor clustering mechanism rather than requirement for a specific homodimeric conformation. The potent agonistic activity attained with these covalent NK1 dimers opens up a myriad of applications in regenerative medicine where the instability and poor expression yield of recombinant HGF have been a severe limitation.

In summary, we used a combination of rational and combinatorial engineering methods to develop receptor agonists and antagonists with therapeutic potential. Because ligand-receptor interactions are important regulators of signaling pathways involved in health and disease, similar protein engineering
strategies have great promise for application to other therapeutically relevant systems.

Methods

Additional information is provided in SI Materials and Methods.

Library Generation. NK1 I1 consists of amino acids Tyr28 to Glu109 of HGF isoform 1 (GenBank Accession ID M1 55000932). NK1 I3 consists of amino acids Tyr28 to Glu205 of HGF isoform 3 (GenBank Accession ID NP_00100932). DNA encoding the open reading frame of NK1 I1 or NK1 I3 was cloned into the pTMY-HA yeast display plasmid using Nhel and MluI restriction sites. The M1 library was generated by error-prone PCR and homologous recombination as previously described (26) using primers with 50 bp of sequence homology to pTMY-HA and equal amounts of NK1 I1 and NK1 I3 plasmid DNA templates. A range of mutation frequencies were obtained using low-fidelity Taq polymerase (Invitrogen) and varying amounts of the nucleotide analogs 8-oxo-dGTP and dTPP (Trilink Biotech) (37) in separate PCR reactions consisting of five cycles (200 μM analogs), 10 cycles (2 μM or 20 μM analogs), or 20 cycles (2 μM analogs). PCR products were amplified in the absence of analogs and a total of 80 μg of mutated cDNA insert and 8 μg restriction-digested pTMY-HA backbone was transformed into EBY100 yeast by electroporation (38). A library of approximately 3 x 10^10 transformants was obtained, as estimated by plating serial dilutions and colony counting. Sequence analysis of randomly selected clones revealed an equal distribution based on NK1 I1 and I3 templates, a mutation rate of 0.2–4.7% for individual mutants, and an average mutation rate of 1.3% for 6,400 sequenced base pairs. The M2 library was created by error-prone PCR using 100 ng DNA from mutants, and an average mutation rate of 1.3% for 6,400 sequenced base pairs. Additional mutations were randomly introduced using nucleotide analogs as described above. A total of 75 μg of mutated cDNA insert and 7.5 μg restriction-digested pTMY-HA backbone was electroporated into EBY100 yeast to generate a library of approximately 8 x 10^10 transformants.

Library Screening. Met-Fc protein was conjugated to Alexa 488 (Met-Fc A488) according to the manufacturer’s instructions except the labeling time was shortened to 10 min to minimize the dye:protein conjugation ratio. Various concentrations of Met-Fc A488 were incubated with yeast-displayed libraries in 1× PBS pH 7.4 containing 1 mg/mL bovine serum albumin (PBS/BSA) for 3 h at room temperature. For the final hour, the anti-HA antibody 6E2 (Cell Signalling Technology) was added at a final dilution of 1:20. Cells were pelleted by centrifugation, washed with 1 mL ice-cold PBS/BSA, and resuspended in PBS/BSA containing a 1:20 dilution of R-Phycocerythrin conjugated goat anti-mouse secondary antibody (Sigma P9670) for 25 min on ice. Cells were washed as above and sorted using a Vantage SE flow cytometer (Stanford FACs Core Facility) and CellQuest software (Becton Dickinson Biosciences). Collected yeast cells were amplified in SD-CAA media pH 4.5 and induced for expression in 5× CAA media at 20, 30, or 37 °C, as appropriate, and additional rounds of FACS were used to obtain enriched pools of mutants. The first round of FACS for M1 and M2 libraries consisted of three separate sorts totaling approximately 1 x 10^5 sorted cells, whereas subsequent sort rounds analyzed at least 10x the number of yeast collected in the previous round to ensure sufficient sampling of remaining library diversity. Sort stringency was increased by decreasing the concentration of Met-Fc A488 (for affinity screening) or increasing induction temperature to 37 °C (for stability screening). Plasmid DNA was recovered from yeast cultures using a Zymoprep kit (Zymo Research) and transformed into XL-1 blue supercompetent Escherichia coli cells (Stratagene) for plasmid miniprep. DNA sequencing was performed by Elii Biopharmaceuticals.

Acknowledgments. We thank Patrick Ma (Case Western Reserve, Cleveland, OH) for BaF3 and BaF3-Met cells, George Vande Woude (Van Andel Institute, Grand Rapids, MI) for M1007 cells, and Birgitte J. Wermuth (University of California, San Diego, CA) for HGF cDNA. This work was funded by National Institutes of Health (NIH) RCI R21 CA131706 and R01 CA151706. D.S.J. was supported by NIH training Grant 5T32 GM008412-15S1 and a Siebel Graduate Fellowship.

Supporting Information

Jones et al. 10.1073/pnas.1102561108

SI Materials and Methods.

Media and Reagents. Yeast extract/pptone/dextrose (YPD) media contained 20 g/L glucose, 20 g/L peptone, and 10 g/L yeast extract. Synthetic dextrose medium with casamino acids (SD-CAA) contained 20 g/L glucose, 6.7 g/L yeast nitrogen base without amino acids, 5.4 g/L Na₃HPO₄, 8.6 g/L NaH₂PO₄ · H₂O, and 5 g/L Bacto casamino acids; SG-CAA media was identical, except glucose was substituted with 20 g/L galactose. SD-CAA pH 4.5 media was identical to SD-CAA, except phosphates were replaced with 13.7 g/L sodium citrate dihydrate, 8.4 g/L citric acid anhydrous, and adjusted to pH 4.5. BMMY and BMGY media and RDB plates for Pichia pastoris strain GS115 were prepared as described in the Multi-Copy Pichia Expression Kit (Invitrogen K1750-01). BMMY media was supplemented with 10 g/L Bacto casamino acids to improve yield with scale-up. Alexa 488 labeling kits (A-20181) were purchased from Invitrogen, and recombinant Met-Fc (358-MT-100-CF) was purchased from R&D Systems. PBS contained sodium phosphate pH 7.4, 137 mM sodium chloride, and 2.7 mM potassium chloride, unless otherwise indicated. BaF3 and BaF3-Met cells (1) were provided by Patrick Ma (Case Western Reserve University, Cleveland, OH). BaF3 complete growth media consisted of RPMI medium 1640 with Glutamax (Invitrogen) supplemented with 10% fetal bovine serum (ATCC), 1% penicillin-streptomycin, and 0.5 ng/mL recombinant mouse interleukin-3 (R&D Systems). Media for BaF3-Met cells was further supplemented with 1 mg/mL Geneticin (Invitrogen) to maintain selection.

Protein Expression and Purification. NK1 proteins, comprising the N terminal (N) and first Kringle domain (K1) of HGF, were recombinantly expressed using the Invitrogen Multi-Copy Pichia Expression Kit. DNA encoding the protein of interest was cloned into the pPIC9K plasmid using SnaBI and NotI sites. An N-terminal FLAG epitope tag (DYKDDDDK) was inserted between SnaBI and AvrII sites, and a C-terminal hexahistidine tag was inserted between MluI and NotI sites. Plasmid DNA (approximately 20 μg) was linearized by digestion with SacI and electroporated into the P. pastoris yeast strain GS115. Transformed yeast cells were allowed to recover on RDB plates for 2 d at 30 °C and were then selected for growth on YPD plates containing 4 mg/mL Geneticin. Multiple Geneticin-resistant colonies were screened for expression by growth in BMGY and subsequent induction in BMMY for 3 d while maintaining methanol concentration at approximately 0.5%. Protein expression was detected by Western blot analysis of culture supernatants, using an antibody against the FLAG epitope tag (Sigma).

The highest expressing colony for each individual mutant was scaled up for expression in shake flask cultures. Protein expression was induced at 20 °C because wild-type NK1 produced at 30 °C was not properly folded as determined by CD spectroscopy. Expression at 20 °C also improved the yield of wild-type and mutant NK1 proteins. Proteins were purified using metal chelating chromatography followed by size-exclusion chromatography. Briefly, hexahistidine-tagged NK1 proteins were bound to HIS-Select Nickel Affinity Gel (Sigma), washed three times with ice-cold 1x PBS containing 500 mM NaCl (PBS500), and eluted with PBS500 containing 500 mM imidazole. NK1 proteins were further purified by size-exclusion chromatography on a Varian ProStar HPLC system, using a Superdex 75 column (GE Healthcare) and PBS500 running buffer. Purified proteins were concentrated using an Amicon Ultra-4 Centrifugal Filter Unit (Millipore) with a 10-kDa molecular weight cutoff membrane. Expected amino acid composition and protein molecular mass were confirmed by amino acid analysis (University of California, Davis Molecular Structure Facility) and MALDI-TOF mass spectrometry (Stanford Protein and Nucleic Acid Facility), respectively. Protein concentration was determined by UV-visible spectroscopy using the method of Scopes (2). Proteins were stored in 40% glycerol at −20 °C for short-term storage or in 30% glycerol at −80 °C for longer-term storage. Proteins used in subsequent biophysical analyses were stored at −80 °C with 0.01% Tween-20. M2.1, M2.2, and M2.2 point mutants had purification yields of 15–30 mg/L, whereas NK1 and NK1 N127A had purification yields of approximately 3.5 mg/L and 0.5 mg/L, respectively.

Binding Affinity. Wild-type or mutant NK1 proteins (0.05–400 nM) were added to 3–6 x 10⁴ BaF3-Met cells in PBS containing 200 mM NaCl and 1 mg/mL bovine serum albumin (PBS200/BSA), using volumes and cell numbers experimentally determined to avoid ligand depletion. After incubation for 5.5 h at 4 °C, cells were washed with ice-cold PBS200/BSA and resuspended with a 1:100 dilution of R-Phycoerythrin (PE)-conjugated anti-FLAG antibody (Prozyme) in 1x PBS containing 1 mg/mL BSA (PBS/BSA) for 25 min on ice. Cells were washed with ice-cold PBS/BSA and analyzed using a FACSCalibur flow cytometer (Becton Dickinson) and FlowJo software (Treestar, Inc.). Fluorescence data from different days was normalized using Sphereo Rainbow Calibration Particles (Spherotech Inc.). Binding titrations were fit to a four-parameter sigmoidal curve using KaleidaGraph (Synergy Software) to determine the equilibrium binding constant (Kd). For binding to Met expressed on A549 lung carcinoma cells, 300 ng/mL basic fibroblast growth factor (Invitrogen) was added for 2.5 h at 4 °C to block heparan-sulfate proteoglycans (HSPGs) on the cell surface. NK1 proteins were then added to final concentrations of 0.05–400 nM. After an additional 5 h at 4 °C, cells were washed, stained with PE-conjugated anti-FLAG antibody, and analyzed by flow cytometry as described above.

Thermal Stability. To test the thermal stability of yeast-displayed NK1 mutants M2.1 or M2.2, 10⁷ yeast cells were suspended in 50 μL PBS/BSA and incubated at various temperatures for 5 min, and washed with 1 mL ice-cold PBS/BSA. Yeast-displayed proteins were then incubated with 20 nM Met-Fc A488 for 3 h at 4 °C. Mouse anti-HA antibody (Cell Signaling Technologies) was added at a final dilution of 1:20 for the last hour of incubation. Cells were washed and resuspended in a 1:20 dilution of PE-conjugated goat anti-mouse secondary antibody (Sigma) and incubated for 25 min on ice. Cells were washed and analyzed by flow cytometry as described above to determine mean fluorescence values. Secondary structure of NK1 proteins was monitored using a Jasco J-815 circular dichroism spectropolarimeter. Ellipticity of 10 μM NK1 proteins in PBS was monitored from 260 to 190 nm in a quartz cuvette with a 1-mm path length. For thermal denaturation studies, ellipticity was monitored at 208 nm using a 1°C/min scan rate. In both assays, melting curves were fit to a standard two-state unfolding curve using KaleidaGraph software.
Heparin Binding. Heparin binding was characterized by analyzing 25 μg of NK1, M2.1, or M2.2 on a 5 mL HiTrap Heparin HP Column (GE Healthcare) using a Varian ProStar HPLC system. PBS running buffer (containing 300 mM NaCl) was used to load proteins onto the column. Proteins were eluted from the column at 0.25 mL/min with a linear gradient from 0.3–2 M NaCl over 100 min. Elution peak retention time was calculated using Galaxy software.

Size-Exclusion Chromatography. To test the effects of reduced salt concentrations, 30–40 μg of NK1, M2.1, or M2.2 was analyzed by size-exclusion chromatography on a Superdex 75 column (GE Healthcare) using PBS running buffer (with 137 mM NaCl) at a flow rate of 0.4 mL/min.

Cellular Met Activation Assays. Met activation assays were performed following published protocols (3). For the MDCK cell scatter assay, 1,500 cells/well were seeded into 96-well plates in 100 μL of complete growth media and incubated at 37 °C, 5% CO₂. After 24 h, media was removed by aspiration and replaced with media containing HGF or NK1 proteins at a concentration of 0.1 or 100 nM, respectively. In some experiments lovenox heparin (Sanofi-Aventis) was used at a concentration of 2 μM or at a 2:1 molar ratio of heparin:NK1. After 24 h, cells were fixed and stained with 0.5% crystal violet in 50% ethanol for 10 min at room temperature, washed with water, and dried in air prior to being photographed. MDCK scatter inhibition assays were performed in a similar manner, except cells were incubated with 250 nM NK1 mutants for 30 min prior to adding HGF at a final concentration of 0.1 nM.

For MDCK uPA assays, 4,000 cells/well were seeded into 96-well plates in 100 μL of complete growth media and incubated at 37 °C, 5% CO₂. After 24 h, media was removed by aspiration and replaced with media containing HGF or NK1 at a concentration of 1 or 100 nM, respectively. After 24 h, cells were washed two times with 200 μL phenol-red free DMEM and incubated with 200 μL reaction buffer containing 50% (vol/vol) of 0.05 units/mL plasminogen (Roche Applied Science), 40% (vol/vol) 50 mM Tris pH 8.0, 10% (vol/vol) and 3 mM chromozym PL (Roche Applied Science) in 100 mM glycine pH 3.5 solution. Plates were incubated for 4 h at 37 °C, 5% CO₂ prior to measuring absorbance at 405 nm using an Infinite M1000 microplate reader (Tecan Group Ltd.).


Fig. S1. (A) Domain structure of HGF. Black arrow indicates site for cleavage into α/β chains. NK1 comprises the N-terminal (N) and first Kringle (K1) domains. SPH: serine protease homology domain. (B) Schematic of the open reading frame of the pTMY-HA yeast display plasmid. Protein of interest is flanked by hemagglutinin (HA) and cmyc epitope tags. A flexible (Gly4Ser3)3 linker (link) is present between the protein of interest and the cmyc epitope tag. These components are genetically fused to the yeast Aga2p mating protein. (C) Schematic of yeast display system. Aga2p is disulfide bonded to the yeast Aga1p protein, which in turn is covalently linked to the yeast cell wall. Display levels of NK1 proteins were measured using a mouse anti-HA primary antibody and goat anti-mouse PE secondary antibody. Binding was measured using a recombinant Fc fusion of the Met receptor extracellular domain, which was fluorescently labeled with Alexa 488 (Met-Fc A488). (D) Yeast-displayed wild-type NK1 I1 binds nonspecifically, as fluorescence levels were unchanged with the addition of 20 or 200 nM Met-Fc A488. In addition, binding levels were similar both in the absence (Top) or presence (Bottom) of 2 μM soluble heparin. (E) The low binding levels of yeast-displayed wild-type NK1 to Met-Fc A488 was unaffected after heating the yeast to various temperatures (which denatures cell-surface proteins), further indicating that binding is nonspecific.
Fig. S2. Heparin binding and surface electrostatics of wild-type and mutant NK1 proteins. (A) Binding to heparin on a HiTrap heparin HP column. Proteins were eluted with a linear gradient of NaCl. Wild-type NK1, M2.1, and M2.2 eluted at salt concentrations of 1.7, 1.2, and 1.3 M NaCl, respectively. (B and C) Surface vacuum electrostatics of the N-domain heparin-binding patch in (B) wild-type NK1 or (C) mutant M2.2. The K62E mutation neutralizes some of the surface potential in the N-domain electropositive cluster. (D and E) Surface vacuum electrostatics of the heparin-binding region in the K1 domain of (D) wild-type NK1 or (E) mutant M2.2. The K170E mutation neutralizes some of the surface potential at the center of the electropositive cluster, whereas K132N neutralizes some of the surface potential on the periphery. Protomer of interest (protomer A) is shown as electrostatic surface model, and ribbon model of protomer B is shown in light gray for reference. N or K1 domains of the respective protomers are noted. Structural models were made using PyMol and PDB ID 1NK1.

Fig. S3. Cellular Met receptor binding assays. (A) The Met antibody AF276, wild-type NK1, and mutant M2.2 all bind to BaF3 cells stably transfected with Met (BaF3-Met), but not to untransfected BaF3 cells. (B) Specificity of binding to the Met receptor in BaF3-Met cells was confirmed by blocking binding with the neutralizing antibody AF276. (C) Preincubation of A549 cells with FGF-2 (FGF+) diminishes, but not completely abolishes, the M2.2 binding signal compared to mock preincubation (FGF−), while not affecting the AF276 binding signal. (D) Specificity of M2.2 to Met and HSPGs on the A549 cell surface was confirmed by complete inhibition of binding through preincubation with both FGF-2 and AF276 (AF276+) compared to partial inhibition through preincubation with only FGF-2 (AF276−). Binding of M2.2 was diminished, but not completely abolished by preincubation with AF276 or FGF-2 alone. Dashed line and negative controls indicate fluorescence levels of cells labeled with secondary antibody only. Representative data from experiments performed on separate days.
Fig. S4. Met receptor binding of wild-type and mutant NK1 proteins. Increasing concentrations of NK1, M2.1, or M2.2 proteins, containing an N-terminal FLAG epitope tag, were titrated against (A) BaF3-Met or (B) A549 cells. For A549 binding, preincubation with 300 ng/ml bFGF was used to block cell-surface HSPGs. Binding of NK1 proteins to cell-surface Met receptor was detected by flow cytometry using a PE-conjugated anti-FLAG antibody. *K*D values are shown ± standard deviation. Representative data of experiments performed on three separate days.

Fig. S5. (A) Thermal stability of yeast-displayed M2.1 and M2.2. Wild-type NK1 was not functionally expressed on the yeast cell surface, so stability could not be assessed in this manner. (B) Far-UV circular dichroism scans of wild-type and NK1 mutants at 20 °C. Scan of wild-type NK1 at 80 °C is also shown. The negative peak at 208 nm owes largely to the β-sheet structural element. The positive peak at 236 nm is attributed to folded aromatic side chains with possible contribution from disulfide bonds (1, 2). (C and D) Correlation of recombinant expression yield to (C) yeast surface expression level or to (D) percentage of yeast that express protein on their surface after induction at 20, 30, or 37 °C. Filled symbols: wild-type NK1 and NK1 N127A; open symbols: engineered NK1 mutants. Error bars represent the standard deviation of experiments performed in triplicate.

Fig. S6. Agonistic activity of wild-type and mutant NK1 proteins. (A) MDCK cell scatter stimulated by HGF (0.1 nM) or NK1 proteins (100 nM) in the absence of heparin. (B) Activity of urokinase-type plasminogen activator (uPA) induced by HGF (1 nM) or NK1 proteins (100 nM) without addition of soluble heparin. (C) MDCK cell scatter assay, testing a range of concentrations of wild-type NK1, M2.2 D127N, or M2.2 D127K in the absence (Top) or presence (Bottom) of 2 μM soluble heparin. Images are representative of experiments performed on separate days. Scale bar: 500 μm.
Fig. S7. Antagonistic activity of wild-type and mutant NK1 proteins. Inhibition of HGF-induced MDCK cell scatter (0.1 nM HGF) by NK1 mutants (250 nM) in the absence (Top) or presence (Bottom) of 2 μM soluble heparin. NK1 N127A is a previously reported antagonist based on wild-type NK1 (1). Images are representative of experiments performed on separate days. Scale bar: 500 μm.

Fig. S8. Covalent NK1 homodimers. (A) Structure of NK1 homodimer (PDB ID code 1NK1) with one protomer in black and one protomer in gray showing the close proximity of respective N termini (in red). (B) FPLC trace of Ni-NTA-purified cystine dimer M2.2 D127N (cdD127N), showing a mixture of dimeric and monomeric species. Absorbance was monitored 280 nm. The dimer peak was collected and used for subsequent cellular activity assays. (C) 12% SDS-PAGE with tris-glycine running buffer of purified cdD127N and cdD127K proteins under nonreduced and reduced conditions, showing the presence of NK1 dimers that dissociate into monomers upon reduction of disulfide bonds. (D) MDCK cell scatter induced by the M2.2 D127N monomer (without the free cysteine residue) or the cystine dimer proteins in the absence of heparin. cdD127N: cystine dimer M2.2 D127N; cdD127K: cystine dimer M2.2 D127K. Agonistic activity of full-length HGF is shown for comparison. Images are representative of experiments performed on separate days. Scale bar: 500 μm.

Table S1. Summary of thermal stability ($T_m$) of wild-type and mutant NK1 proteins as determined by CD temperature melts

<table>
<thead>
<tr>
<th>Protein</th>
<th>$T_m \pm \text{SD, } ^\circ\text{C}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NK1</td>
<td>50.7 ± 0.2</td>
</tr>
<tr>
<td>NK1 N127A</td>
<td>47.9 ± 0.7</td>
</tr>
<tr>
<td>M2.1</td>
<td>63.9 ± 0.5</td>
</tr>
<tr>
<td>M2.2</td>
<td>69 ± 1</td>
</tr>
<tr>
<td>M2.2 D127N</td>
<td>65.5 ± 0.5</td>
</tr>
<tr>
<td>M2.2 D127A</td>
<td>63.7 ± 0.1</td>
</tr>
<tr>
<td>M2.2 D127K</td>
<td>62.5 ± 0.1</td>
</tr>
<tr>
<td>M2.2 D127R</td>
<td>62.3 ± 0.5</td>
</tr>
<tr>
<td></td>
<td>N domain</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>% ID†</td>
<td>62</td>
</tr>
<tr>
<td>huHGF</td>
<td>100</td>
</tr>
<tr>
<td>M2.1</td>
<td>NA</td>
</tr>
<tr>
<td>M2.2</td>
<td>NA</td>
</tr>
<tr>
<td>huMSP</td>
<td>38</td>
</tr>
<tr>
<td>Cat</td>
<td>94</td>
</tr>
<tr>
<td>Dog</td>
<td>92</td>
</tr>
<tr>
<td>Bovine</td>
<td>94</td>
</tr>
<tr>
<td>Mouse</td>
<td>90</td>
</tr>
<tr>
<td>Chicken</td>
<td>65</td>
</tr>
<tr>
<td>huHGF K2¶</td>
<td>44</td>
</tr>
<tr>
<td>huHGF K3¶</td>
<td>52</td>
</tr>
<tr>
<td>huHGF K4§</td>
<td>37</td>
</tr>
</tbody>
</table>

Blank spaces indicate same residue as human HGF (huHGF). HGF isoform 1 numbering scheme is used.
*Linker region connecting N and K1 domains of NK1.
†Percent sequence identity with the huHGF NK1 region, or with the huHGF K1 domain for huHGF K2, K3, or K4.
‡There is no human macrophage stimulating protein (huMSP) residue corresponding to huHGF K170.
¶No N domain or linker region.
§Alignment of huHGF Kringle domain 2, 3, or 4 to huHGF Kringle domain 1.