

BACTERIAL LON PROTEASE DOMAINS AND THEIR USE IN THERAPY

TECHNICAL FIELD

The invention relates to new compositions, particularly compositions comprising one or more domains of bacterial Lon protease or related peptides. The compositions are useful for the treatment of various diseases, such as infections, particularly bacterial infections, and cancers.

BACKGROUND

The MYC transcription factors are important therapeutic targets. They remain "undruggable" due to the lack of specific MYC inhibitors.

Transcriptional control is essential to maintain normal tissue functions and optimize the defense against common diseases such as infections or cancer. The MYC transcription factors are commonly regarded as "master gene regulators," which control the expression of >15% of the human genome. It has been stated that all crucial aspects of cell proliferation, cell growth and tumorigenesis are positively regulated by C-MYC and in the right amount, C-MYC is essential for normal cell function, but its deregulation is extremely dangerous. Indeed, MYC deregulation has been proposed to drive tumor development in >70% of human cancers, such as colorectal carcinoma, breast cancer, prostate cancer and liver cancer. Finding disease-specific MYC inhibitors has been problematic, however, and despite extensive efforts, drugs targeting MYC are not available for clinical use. The lack of specific MYC inhibitors has also made it difficult to define how MYC deregulation affects diseases states other than cancer ^{6,7}.

The recently discovered Lon protease produced by uropathogenic *Escherichia coli* degrades C-MYC, attenuates C-MYC-dependent gene expression and delays tumor progression. This study investigated the MYC response to bacterial kidney infection, and the protective potential of LON treatment for this indication. The recombinant Lon protease (rLon) was identified as a potent pan-MYC inhibitor, protecting susceptible *Irf3*^{-/-} mice from renal tissue pathology and accelerating bacterial clearance. A similar effect was detected in intestinal cancer tissues from rLon treated *Apc*^{Min/+} mice, where the overexpression of C-MYC, N-MYC and L-MYC was inhibited by rLon treatment. By peptide mapping, the LON peptidase domain was shown to degrade C-MYC, N-MYC and L-MYC in human kidney cells and the N-terminal domain (NTD) was identified as a transcriptional L-MYC inhibitor. The results identify rLon and related peptides as a remarkably efficient pan-MYC inhibitor with a dual mechanism of action and marginal effects on healthy tissues, suggesting a promising therapeutic potential.

35

SUMMARY OF THE INVENTION

The invention provides a peptide comprising an amino acid sequence having substantial identity to the amino acid sequence of one or more domains of Lon protease or a functional fragment or variant thereof.

Lon protease is an ATP-dependent serine peptidase. Examples of Lon protease include those secreted by bacteria, such as *E. coli*. The peptide of the invention may comprise an amino acid sequence having substantial identity to the amino acid sequence of one or more domains of a Lon protease that is of bacterial origin. By this, we mean that these molecules have a structure that is of bacterial origin, and therefore covers synthetic reproductions of molecules of bacterial origin. In the case of proteins, for example, the proteins could be biosynthesized by the bacteria, biosynthesized recombinantly as exogenous molecules in an artificial expression system or could be artificial synthetic proteins that match the bacterial protein sequence. Inhibitors may be obtainable from a wide range of bacterial species, including for example, *Streptococcus*, *Pseudomonas*, *Shigella*, *Campylobacter*, and *Salmonella* species. In a particular embodiment the bacteria is a uropathogenic bacteria, preferably *E. coli*, more preferably *E. coli* 536 or CFT073.

When the domain is the N-terminal domain, the Lon protease is generally prokaryotic in origin, particularly bacterial origin.

Other non-limiting examples of Lon protease are provided in the following table.

Description	Organism	Accession
lon protease	<i>Escherichia coli</i>	NP_308520.1
endopeptidase La	<i>Shewanella oneidensis</i>	WP_011071918.1
Lon protease	<i>Pseudomonas aeruginosa</i>	NP_250494.1
endopeptidase La	<i>Neisseria meningitidis</i>	WP_002229478.1
class III heat-shock ATP-dependent LonA protease	<i>Bacillus subtilis</i>	NP_390698.1
endopeptidase La	<i>Clostridioides difficile</i>	WP_021364834.1
endopeptidase La	<i>Deinococcus radiodurans</i>	WP_027479826.1
endopeptidase La	<i>Thermotoga maritima</i>	WP_004082110.1

endopeptidase La	<i>Streptomyces</i>	WP_011030152.1
lon protease 2	<i>Arabidopsis thaliana</i>	NP_568675.1
lon protease homolog 2, peroxisomal	<i>Glycine max</i>	XP_003524333.1
Lon protease, isoform A	<i>Drosophila melanogaster</i>	NP_649133.1
lon protease homolog 2, peroxisomal isoform 1	<i>Mus musculus</i>	NP_080103.1
lon protease homolog, mitochondrial	<i>Danio rerio</i>	XP_691712.4
Lon protease	<i>Schizosaccharomyces pombe</i>	NP_593035.1
Lon protease homolog, mitochondrial	<i>Caenorhabditis elegans</i>	NP_492796.1
ATP-dependent Lon protease PIM1	<i>Saccharomyces cerevisiae</i>	NP_009531.1
peptidase S16, Lon protease family protein	<i>Dictyostelium discoideum</i>	XP_642688.1
lon protease homolog 2, peroxisomal isoform X2	<i>Homo sapiens</i>	XP_016879245.1

Table 1.

The amino acid sequence of a Lon protease from an example *E. coli* strain is as follows:

MNPERSERIEIPVLPLRDVVVYPHVIPLFVGREKSIRCLEAAMDHDKKIMLVAQKEASTDEPGVNDLFT
 5 VGTVASILQMLKLPDGTVKVLVEGLQRARISALSDNGEHFSAKAEYLESPTIDEREQEVLVRTAISQFEG
 YIKLNKKIPPEVLTSLNSIDDPARLADTIAAHMPLKLADKQSVLEMSDVNERLEYLMAMMESEIDLLQVEK
 RIRNRVKKQMEKSQREYYLNEQMKAIQKELGEMDDAPDENEALKRKIDAAMPKEAKEKAEELQKLK
 MMSPMSAEATVVRGYIDWMVQVPWNARSKVKKDLRQAQEILDTDHYGLERVKDRILEYLAVQSRVNKI
 10 KGPILCLVGPPGVGKTSLGQSIKATGRKYVRMALGGVRDEAEIRGHRRTYIGSMPGKLIQKMAKVGVK
 NPLFLLDEIDKMSSDMRGDPASALLEVLDPENQVAFSDHYLEVVDYDLSVDMFVATSNMNIAPLLDRM
 EVIRLSGYTEDEKLNIAKRHLLPKQIERNALKKGELTVDDSAIIGIIRYYTREAGVRGLEREISKLCRKAVK
 QLLLDKSLKHIEINGDNLHDYLGVRFDYGRADNENRVGQVTGLAWTEVGGDLLTIETACVPGKGKLTYY

TGSLGEVMQESIQAALTVVRRARAEKLGINPDFYEKRDIVHVHVEGATPKDGPSAGIAMCTALVSCLTGN
PVRADVAMTGEITLRGQVLPIGGLKEKLLAAHRGGIKTVLIPFENKRDLEEIPDNVIADLDIHPVKRIEEVL
TLALQNEPSGMQVVTAK (SEQ. ID No. 1).

Lon protease comprises an N-terminal domain, an AAA+ domain and a C-terminal
5 proteolytic domain or peptidase. One skilled in the art can determine the domains in any
Lon protease using methods described in the Examples section.

The N-terminal domain generally comprises a region of alpha helices. In *E. coli* originating
Lon protease, the N-terminal domain is generally found between amino acids 1 and 230,
235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305 or 310. The
10 alpha helices are generally found between amino acids 100, 105, 110, 115, 116, 117, 118,
119, 120, 121, 122, 123, 124 or 125 and 230, 235, 240, 241, 242, 243, 244, 245, 246,
247, 248, 249, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, 300, 305 or 310.

In *E. coli* originating Lon protease, the AAA+ domain is generally found between amino
acids 230, 235, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 255, 260, 265,
15 270, 275, 280, 285, 290, 295, 300, 305 or 310 and 560, 565, 570, 575, 580, 581, 582,
583, 584, 585, 586, 587, 588, 589, 590, 595, 600, 605, or 610.

In *E. coli* originating Lon protease, the C-terminal proteolytic domain is generally found
between amino acids 560, 565, 570, 575, 580, 581, 582, 583, 584, 585, 586, 587, 588,
589, 590, 595, 600, 605, or 610 and 780, 781, 782, 783, 784, 785, 786, 787, 788, 789 or
20 790.

In the peptide of the invention, the Lon protease domain may be the N-terminal domain
either including or excluding, wholly or partly the alpha helices, the AAA+ domain or the C-
terminal domain.

Accordingly, the invention provides a peptide comprising an amino acid sequence having
25 substantial identity to the amino acid sequence of the N-terminal domain of Lon protease or
a functional fragment or variant thereof. The N-terminal domain may comprise the alpha
helix region, or that region may be partly or wholly omitted.

Further, the invention provides a peptide comprising an amino acid sequence having
substantial identity to the amino acid sequence of the AAA+ domain of Lon protease or a
30 functional fragment or variant thereof.

Further still, the invention provides a peptide comprising an amino acid sequence having
substantial identity to the amino acid sequence of the C-terminal proteolytic domain of Lon
protease or a functional fragment or variant thereof.

In an embodiment, the peptide comprises an amino acid sequence having substantial identity to the amino acid sequence of a first Lon protease domain or a functional fragment or variant thereof and an amino acid sequence having substantial identity to the amino acid sequence of a second Lon protease domain or a functional fragment or variant thereof. In other words, part of the amino acid sequence of the peptide has substantial identity to the amino acid sequence of a first Lon protease domain or a functional fragment or variant thereof and another part has substantial identity to the amino acid sequence of a second Lon protease domain or a functional fragment or variant thereof.

In an embodiment, it comprises an amino acid sequence having substantial identity to only one domain, or only two domains of Lon protease or functional fragments or variants thereof. The peptide may comprise other amino acids, or may consist of an amino acid sequence having substantial identity to one or two Lon protease domains.

Optionally, the peptide may comprise an amino acid sequence having substantial identity to the amino acid sequence of the N-terminal domain of Lon protease or a functional fragment or variant thereof and an amino acid sequence having substantial identity to the amino acid sequence of another domain, such as the C-terminal proteolytic domain of Lon protease or a functional fragment or variant thereof.

Optionally, the peptide may comprise an amino acid sequence having substantial identity to the amino acid sequence of the C-terminal proteolytic of Lon protease or a functional fragment or variant thereof and an amino acid sequence having substantial identity to the amino acid sequence of another domain, such as the N-terminal domain of Lon protease or a functional fragment or variant thereof.

In an embodiment, the peptide comprises an amino acid sequence having substantial identity to the amino acid sequence of a first Lon protease domain or a functional fragment or variant thereof and an amino acid sequence having substantial identity to the amino acid sequence of a second Lon protease domain or a functional fragment or variant thereof. In other words, part of the amino acid sequence of the peptide has substantial identity to the amino acid sequence of a first Lon protease domain or a functional fragment or variant thereof and another part has substantial identity to the amino acid sequence of a second Lon protease domain or a functional fragment or variant thereof.

In an embodiment, it comprises an amino acid sequence having substantial identity to only one domain, or only two domains of Lon protease. The peptide may comprise other regions, or may consist of an amino acid sequence having substantial identity to one or two Lon protease domains.

The peptide does not generally consist of the amino acid sequence of a wild-type Lon protease.

The peptide may, for example, be less than about 780, 770, 760, 750, 740, 730, 720, 710, 700, 690, 680, 670, 660, 650, 640, 630, 620, 610, 600, 590, 580, 570, 560, 550, 540, 530, 520, 510, 500, 490, 480, 470, 460, 450, 440, 430, 420, 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200,
 5 190, 180, 170, 160, 150, 140, 130 or 120 amino acids in length.

When the peptide comprises an amino acid sequence having substantial identity to the amino acid sequence of the N-terminal domain of Lon protease, or a functional fragment or variant thereof, it may optionally be more than 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 236, 238, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 252 or
 10 254 amino acids in length. It may also, or alternatively, optionally be less than 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 249, 248, 247, 246, 245, 244, 243, 242, 241, 240, 230, 220, 210, 200, 190, 180, 170, 160, 150, 140, 130, 129, 128, 127, 126, 125, 124, 123, 122, 121, 120, 119, 118, 117, 116 or 115 amino acids in length.

When the peptide comprises an amino acid sequence having substantial identity to the
 15 amino acid sequence of the C-terminal proteolytic domain of Lon protease, or a functional fragment or variant thereof, it may optionally be more than 180, 185, 190, 195, 200, 205, 210, 215 or 220 amino acids in length. It may also, or alternatively, optionally be less than 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 170, 160, 150, 140, 130 or 120 amino acids in length.

20 When the peptide comprises an amino acid sequence having substantial identity to the amino acid sequence of the N-terminal domain of Lon protease, or a functional fragment or variant thereof, and an amino acid sequence having substantial identity to the amino acid sequence of the peptidase domain of Lon protease, or a functional fragment or variant thereof, those sequences may be contiguous, or may be separated by other amino acids.

25 In an embodiment, the N-terminal domain has one of the following sequences:

MNPERSERIEIPVLPLRDVVVYPHMYIPLFVGREKSIRCLEAAMDHDKKIMLVAQKEASTDEPGVNDLFT
 VGTVASILQMLKLPDGTVKVLVEGLQRARISALSDNGEHFSAKAEYLE (SEQ. ID No. 2 (without
 alpha helices)) and

MNPERSERIEIPVLPLRDVVVYPHMYIPLFVGREKSIRCLEAAMDHDKKIMLVAQKEASTDEPGVNDLFT
 30 VGTVASILQMLKLPDGTVKVLVEGLQRARISALSDNGEHFSAKAEYLESPTIDEREQEVLRVTAISQFEG
 YIKLNKKIPPEVLTSLNSIDDPARLADTIAAHMPLKLADKQSVLEMSDVNERLEYLMAMMESEIDLLQVEK
 RIRNRVKKQMEKSQREYYLNEQMKAIQKELGEMD (SEQ. ID No. 3 (with alpha helices)).

Accordingly, the invention provides a peptide comprising an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID No. 2 or 3 or a fragment or

variant thereof. Identity, fragment and variant are as defined in relation to other embodiments or aspects of the invention.

In an embodiment, the AAA+ domain comprises the following amino acid sequence:

DAPDENEALKRKIDAAKMPKEAKEAEAEELQKLKMMSPMSAEATVVRGYIDWMVQVPWNARSKVKKD
 5 LRQAQEILDTDHYGLERVKDRILEYLAVQSRVNKIKGPILCLVPPGVGKTSLGQSIKATGRKYVRMAL
 GGVREAEIRGHRRTYIGSMPGKLIQKMAKVGKVNPLFLLDEIDKMSSDMRGDPASALLEVLDPENVA
 FSDHYLEVDYDLSDVMFVATSNMNIAPLLDRMEVIRLSGYTEDEKLNIAKRHLLPKQIERNALKKGEL
 TVDDSAIIGIIRYYTREAGVRGLEREISKLCRKAVKQLLLDKSLKHIEINGDNLHDYLGVRFDYGRADNE
 NRVGQVTG (SEQ ID No. 4).

- 10 Accordingly, the invention provides a peptide comprising an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID No. 4 or a fragment or variant thereof.

In an embodiment, the C-terminal domain comprises the following amino acid sequence:

LAWTEVGDLTIETACVPGKGKLTGTSLGEVMQESIQAALTVVRARAELGINPDFYEKRDHVVHVE
 15 GATPKDGPSAGIAMCTALVSCLTGNPVRADVAMTGEITLRGQVLPIGGLKEKLLAAHRRGGIKTVLIPFEN
 KRDLEEIPDNVIADLDIHPVKRIEEVLTALQNEPSGMQVVTAK (SEQ ID No. 5).

Accordingly, the invention provides a peptide comprising an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID No. 5 or a fragment or variant thereof.

- 20 Optionally, the peptide may comprise an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID. 2 or 3 and 5 or a fragment or variant thereof.

The term substantial identity means the sequence is similar to the sequence referenced. For example, substantial identity may mean about or at least about 60%, 65%, 70%, 72%, 75%, 78%, 80%, 82%, 85%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%,
 25 or 99% identity or homology.

Accordingly, the peptide may comprise an amino acid sequence having about or at least about 60%, 65%, 70%, 72%, 75%, 78%, 80%, 82%, 85%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity or homology to the amino acid sequence of the N-terminal domain of Lon protease, such as to SEQ ID No. 2 or 3, or to a functional
 30 fragment or variant thereof. The N-terminal domain may comprise the alpha helices, or they may be wholly or partly omitted.

Alternatively, the peptide may comprise an amino acid sequence having about or at least about 60%, 65%, 70%, 72%, 75%, 78%, 80%, 82%, 85%, 88%, 90%, 91%, 92%, 93%,

94%, 95%, 96%, 97%, 98%, or 99% identity or homology to the amino acid sequence of the C-terminal proteolytic domain, such as to SEQ ID No. 5 of Lon protease or a functional fragment or variant thereof.

Alternatively, the peptide may comprise an amino acid sequence having about or at least
5 about 60%, 65%, 70%, 72%, 75%, 78%, 80%, 82%, 85%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity or homology to the amino acid sequence of the AAA+ domain of Lon protease, such as to SEQ ID No. 4, or to a functional fragment or variant thereof.

Alternatively, the peptide may comprise an amino acid sequence having about or at least
10 about 60%, 65%, 70%, 72%, 75%, 78%, 80%, 82%, 85%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence of the N-terminal domain of Lon protease, such as SEQ ID No. 2 or 3, and an amino acid sequence having about or at least about 60%, 65%, 70%, 72%, 75%, 78%, 80%, 82%, 85%, 88%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to the amino acid sequence
15 of the C-terminal proteolytic domain, such as SEQ ID No. 5 of Lon protease.

A variant will preferably have at least about 40% homology to the amino acid sequence of the peptide over its entire length. More preferably, the variant may have at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least
20 about 90%, at least about 95%, at least about 97% or at least about 99% homology over the entire sequence.

A variant will preferably have at least about 40% identity to the amino acid sequence of the peptide over its entire length. More preferably, the variant may have at least about 45%, at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least
25 about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 97% or at least about 99% identity over the entire sequence.

There may be at least about 80%, for example at least about 85%, about 90%, about 95%, about 98% or about 99%, homology or identity over a stretch of 75 or more, for example
30 80, 90, 100, 125, 150, 175 or 200 or more, contiguous amino acids of the peptide ("hard homology").

Standard methods in the art may be used to determine homology or identity. For example, the UWGCG Package provides the BESTFIT program which can be used to calculate homology, for example used on its default settings (Devereux et al (1984) Nucleic Acids
35 Research 12, p387-395). The PILEUP and BLAST algorithms can be used to calculate homology or line up sequences (such as identifying equivalent residues or corresponding

sequences (typically on their default settings)), for example as described in Altschul S. F. (1993) J Mol Evol 36:290-300; Altschul, S.F et al (1990) J Mol Biol 215:403-10. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>).

- 5 The variant may comprise one or more amino acid substitutions, for example at least about 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 or more substitutions. The one or more substitutions are preferably one or more conservative substitutions. Conservative substitutions replace amino acids with other amino acids of similar chemical structure, similar chemical properties, or similar side-chain volume. The amino acids introduced may
- 10 have similar polarity, hydrophilicity, hydrophobicity, basicity, acidity, neutrality, or charge to the amino acids they replace. Alternatively, the conservative substitution may introduce another amino acid that is aromatic or aliphatic in the place of a pre-existing aromatic or aliphatic amino acid. Conservative amino acid changes are well-known in the art and may be selected in accordance with the properties of the 20 main amino acids as defined in Table
- 15 2 below. Where amino acids have similar polarity, this can also be determined by reference to the hydropathy scale for amino acid side chains in Table 3.

Table 2 – Chemical properties of amino acids

Ala	aliphatic, hydrophobic, neutral	Met	hydrophobic, neutral
Cys	polar, hydrophobic, neutral	Asn	polar, hydrophilic, neutral
Asp	polar, hydrophilic, charged (-)	Pro	hydrophobic, neutral
Glu	polar, hydrophilic, charged (-)	Gln	polar, hydrophilic, neutral
Phe	aromatic, hydrophobic, neutral	Arg	polar, hydrophilic, charged (+)
Gly	aliphatic, neutral	Ser	polar, hydrophilic, neutral
His	aromatic, polar, hydrophilic, charged (+)	Thr	polar, hydrophilic, neutral
Ile	aliphatic, hydrophobic, neutral	Val	aliphatic, hydrophobic, neutral
Lys	polar, hydrophilic, charged(+)	Trp	aromatic, hydrophobic, neutral
Leu	aliphatic, hydrophobic, neutral	Tyr	aromatic, polar, hydrophobic

Table 3 - Hydropathy scale

	Side Chain	Hydropathy
5	Ile	4.5
	Val	4.2
	Leu	3.8
	Phe	2.8
10	Cys	2.5
	Met	1.9
	Ala	1.8
	Gly	-0.4
	Thr	-0.7
15	Ser	-0.8
	Trp	-0.9
	Tyr	-1.3
	Pro	-1.6
	His	-3.2
20	Glu	-3.5
	Gln	-3.5
	Asp	-3.5
	Asn	-3.5
	Lys	-3.9
25	Arg	-4.5

Methods for introducing or substituting naturally-occurring amino acids are well known in the art. For instance, methionine (M) may be substituted with arginine (R) by replacing the codon for methionine (ATG) with a codon for arginine (CGT) at the relevant position in a polynucleotide encoding the variant. Methods for introducing or substituting non-naturally-occurring amino acids are also well known in the art. For instance, non-naturally-occurring amino acids may be introduced by including synthetic aminoacyl-tRNAs in the IVTT system used to express the mutant monomer. Alternatively, they may be introduced by expressing the mutant monomer in *E. coli* that are auxotrophic for specific amino acids in the presence of synthetic (i.e. non-naturally-occurring) analogues of those specific amino acids. They may also be produced by naked ligation if the mutant monomer is produced using partial peptide synthesis.

One or more amino acids may additionally be deleted from the peptides described above.
Up to 1, 2, 3, 4, 5, 10, 20 or 30 or more amino acids may be deleted.

Fragments may be at least 50, at least 100, at least 150, at least 200 or at least 250 amino acids in length.

- 5 One or more amino acids may be alternatively or additionally added to the peptides described above. An extension may be provided at the amino terminal or carboxy terminal of the peptide, variant or fragment. The extension may be quite short, for example from 1 to 10 amino acids in length. Alternatively, the extension may be longer, for example up to 50 or 100 amino acids. A carrier peptide or protein may be fused to a peptide of the
10 invention.

- The fragments or variants are generally functional. A functional fragment or variant means a fragment or variant that demonstrates the same function as the peptide of the invention. For example, a functional fragment of the Lon protease N-terminal domain inhibits L-MYC, C-MYC and N-MYC. A functional fragment of the Lon protease C-terminal proteolytic domain
15 degrades C-MYC, N-MYC and L-MYC. Methods for measuring the activity of the wild-type peptide are known in the art and are provided in the examples. A variant may have increased or decreased activity when compared to the wild-type peptide, such as an increase or decrease of about 1%, 2%, 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60% or more.
- 20 The peptide, variant or fragment may be isolated, substantially isolated, purified or substantially purified. A peptide, variant or fragment is isolated or purified if it is completely free of any other components, such as lipids or other peptides. A peptide, variant or fragment is substantially isolated if it is mixed with carriers or diluents which will not interfere with its intended use. For instance, a peptide, variant or fragment is substantially
25 isolated or substantially purified if it is present in a form that comprises less than 10%, less than 5%, less than 2% or less than 1% of other components.

- In all of the discussion herein, the standard one letter codes for amino acids are used. These are as follows: alanine (A), arginine (R), asparagine (N), aspartic acid (D), cysteine (C), glutamic acid (E), glutamine (Q), glycine (G), histidine (H), isoleucine (I), leucine (L),
30 lysine (K), methionine (M), phenylalanine (F), proline (P), serine (S), threonine (T), tryptophan (W), tyrosine (Y) and valine (V). Standard substitution notation is also used, i.e., DXXN means that D at position XX is replaced with N.

- Any of the peptides described herein including variants and fragments may be modified to assist their identification or purification, for example by the addition of histidine residues (a
35 his tag), aspartic acid residues (an asp tag), a streptavidin tag, a flag tag, a SUMO tag, a GST tag or a MBP tag, or by the addition of a signal sequence to promote their secretion

from a cell where the peptide does not naturally contain such a sequence. An alternative to introducing a genetic tag is to chemically react a tag onto a native or engineered position on the peptide. An example of this would be to react a gel-shift reagent to a cysteine engineered on the outside of the peptide.

- 5 Any of the peptides described herein including variants and fragments may be labelled with a revealing label. The revealing label may be any suitable label which allows the peptide to be detected. Suitable labels include, but are not limited to, fluorescent molecules, radioisotopes, e.g., ¹²⁵I, ³⁵S, enzymes, antibodies, antigens, polynucleotides, and ligands such as biotin.
- 10 Any of the peptides described herein including variants and fragments may be made synthetically or by recombinant means. For example, the peptide may be synthesised by in vitro translation and transcription (IVTT). The amino acid sequence of the peptide may be modified to include non-naturally occurring amino acids or to increase the stability of the peptide. When a peptide is produced by synthetic means, such amino acids may be
- 15 introduced during production. Polynucleotide sequences encoding a peptide may be derived and replicated using standard methods in the art. Polynucleotide sequences encoding a peptide may be expressed in a bacterial host cell using standard techniques in the art. This is discussed in more detail below. The peptide may also be altered following either synthetic or recombinant production.
- 20 Peptides may also be produced using D-amino acids. For instance, the peptide may comprise a mixture of L-amino acids and D-amino acids. This is conventional in the art for producing such peptides.

- The peptide may also contain other non-specific modifications as long as they do not interfere with the function of the peptide. A number of non-specific side chain modifications
- 25 are known in the art and may be made to the side chains of the peptide(s). Such modifications include, for example, reductive alkylation of amino acids by reaction with an aldehyde followed by reduction with NaBH₄, amidination with methylacetimidate or acylation with acetic anhydride.

- Peptides may be produced in large scale following purification by any protein liquid
- 30 chromatography system from protein producing organisms or after recombinant expression. Typical protein liquid chromatography systems include FPLC, AKTA systems, the Bio-Cad system, the Bio-Rad BioLogic system and the Gilson HPLC system.

The invention further provides a polynucleotide encoding a peptide according to the invention.

The term polynucleotide refers to a polymer comprising two or more nucleotides. Preferably, the polynucleotide comprises at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides or at least 100 nucleotides. The nucleotides can be naturally occurring or artificial.

- 5 A nucleotide typically contains a nucleobase, a sugar and at least one linking group, such as a phosphate, 2'-O-methyl, 2' methoxy-ethyl, phosphoramidate, methylphosphonate or phosphorothioate group. The nucleobase is typically heterocyclic. Nucleobases include, but are not limited to, purines and pyrimidines and more specifically adenine (A), guanine (G), thymine (T), uracil (U) and cytosine (C). The sugar is typically a pentose sugar. Nucleotide
10 sugars include, but are not limited to, ribose and deoxyribose. The sugar and the nucleobase together form a nucleoside. Preferred nucleosides include, but are not limited to, adenosine, guanosine, 5-methyluridine, uridine, cytidine, deoxyadenosine, deoxyguanosine, thymidine, deoxyuridine and deoxycytidine. The nucleosides may be adenosine, guanosine, uridine and cytidine.
- 15 The nucleotides are typically ribonucleotides or deoxyribonucleotides. The nucleotides may be deoxyribonucleotides. The nucleotides typically contain a monophosphate, diphosphate, or triphosphate. Phosphates may be attached on the 5' or 3' side of a nucleotide.

- Nucleotides include, but are not limited to, adenosine monophosphate (AMP), adenosine diphosphate (ADP), adenosine triphosphate (ATP), guanosine monophosphate (GMP),
20 guanosine diphosphate (GDP), guanosine triphosphate (GTP), thymidine monophosphate (TMP), thymidine diphosphate (TDP), thymidine triphosphate (TTP), uridine monophosphate (UMP), uridine diphosphate (UDP), uridine triphosphate (UTP), cytidine monophosphate (CMP), cytidine diphosphate (CDP), cytidine triphosphate (CTP), 5-methylcytidine monophosphate, 5-methylcytidine diphosphate, 5-methylcytidine triphosphate, 5-
25 hydroxymethylcytidine monophosphate, 5-hydroxymethylcytidine diphosphate, 5-hydroxymethylcytidine triphosphate, cyclic adenosine monophosphate (cAMP), cyclic guanosine monophosphate (cGMP), deoxyadenosine monophosphate (dAMP), deoxyadenosine diphosphate (dADP), deoxyadenosine triphosphate (dATP), deoxyguanosine monophosphate (dGMP), deoxyguanosine diphosphate (dGDP), deoxyguanosine
30 triphosphate (dGTP), deoxythymidine monophosphate (dTMP), deoxythymidine diphosphate (dTDP), deoxythymidine triphosphate (dTTP), deoxyuridine monophosphate (dUMP), deoxyuridine diphosphate (dUDP), deoxyuridine triphosphate (dUTP), deoxycytidine monophosphate (dCMP), deoxycytidine diphosphate (dCDP) and deoxycytidine triphosphate (dCTP), 5-methyl-2'-deoxycytidine monophosphate, 5-methyl-2'-deoxycytidine diphosphate,
35 5-methyl-2'-deoxycytidine triphosphate, 5-hydroxymethyl-2'-deoxycytidine monophosphate, 5-hydroxymethyl-2'-deoxycytidine diphosphate and 5-hydroxymethyl-2'-deoxycytidine triphosphate. The nucleotides may be selected from AMP, UMP, GMP, CMP,

dAMP, dTMP, dGMP or dCMP. In some embodiments, the nucleotides are selected from dAMP, dTMP, dGMP or dCMP.

The nucleotides may contain additional modifications. In particular, suitable modified nucleotides include, but are not limited to, 2'-amino pyrimidines (such as 2'-amino cytidine and 2'-amino uridine), 2'-hydroxyl purines (such as , 2'-fluoro pyrimidines (such as 2'-fluorocytidine and 2'-fluoro uridine), hydroxyl pyrimidines (such as 5'- α -P-borano uridine), 2'-O-methyl nucleotides (such as 2'-O-methyl adenosine, 2'-O-methyl guanosine, 2'-O-methyl cytidine and 2'-O-methyl uridine), 4'-thio pyrimidines (such as 4'-thio uridine and 4'-thio cytidine) and nucleotides have modifications of the nucleobase (such as 5-pentynyl-2'-deoxy uridine, 5-(3-aminopropyl)-uridine and 1,6-diaminohexyl-N-5-carbamoylmethyl uridine).

One or more nucleotides in the polynucleotide may be modified, for instance with a label or a tag. The label may be any suitable label which allows the nucleotides to be detected. Suitable labels include, but are not limited to, fluorescent molecules, radioisotopes, e.g., ^{125}I , ^{35}S , enzymes, antibodies, antigens, other polynucleotides and ligands such as biotin.

The nucleotides in the polynucleotide may be attached to each other in any manner. The nucleotides may be linked by phosphate, 2'-O-methyl, 2' methoxy-ethyl, phosphoramidate, methylphosphonate or phosphorothioate linkages. The nucleotides are typically attached by their sugar and phosphate groups. The nucleotides may be connected via their nucleobases as in pyrimidine dimers.

The polynucleotide may comprise a deoxyribonucleic acid (DNA) or a ribonucleic acid (RNA). Preferably, the polynucleotide comprises DNA. The polynucleotide may be any synthetic nucleic acid known in the art, such as peptide nucleic acid (PNA), glycerol nucleic acid (GNA), threose nucleic acid (TNA), locked nucleic acid (LNA), morpholino nucleic acid or other synthetic polymers with nucleotide side chains.

Substitutions may be used for the practices of codon optimisation and codon wobble, both of which are known to those skilled in the art. Thus, it will be appreciated that codon-optimised and codon-wobbled isolated polynucleotides are also envisaged. In some embodiments, the polynucleotide is codon-optimised for human expression.

The polynucleotide can be produced by de novo solid-phase DNA synthesis or by PCR mutagenesis of an existing sequence. Direct chemical synthesis of polynucleotides can be accomplished by methods known in the art, such as the phosphotriester method of Narang et al., 1979, Meth. Enzymol. 68:90; the phosphodiester method of Brown et al., 1979, Meth. Enzymol. 68:109; the diethylphosphoramidite method of Beaucage et al., 1981, Tetra. Lett., 22:1859; and the solid support method of U.S. Pat. No. 4,458,066. Introducing mutations to a polynucleotide sequence by PCR can be performed as described in, e.g., PCR

Technology: Principles and Applications for DNA Amplification, H. A. Erlich (Ed.), Freeman Press, NY, N.Y., 1992; PCR Protocols: A Guide to Methods and Applications, Innis et al. (Ed.), Academic Press, San Diego, Calif, 1990; Mattila et al., 1991, Nucleic Acids Res. 19:967; and Eckert et al., 1991, PCR Methods and Applications 1:17.

- 5 The invention also provides a vector comprising the polynucleotide of the invention.

The invention also provides a host cell comprising the polynucleotide of the invention or the vector of the invention.

- Various viral and non-viral vectors are known to those skilled in the art. Non-viral vectors include plasmids, episomal vectors, and human artificial chromosomes (see, e.g.,
- 10 Harrington et al., 1997, Nat Genet. 15:345). For example, non-viral vectors useful for expression of the amiRNA and the CAR of the invention in mammalian (e.g., human) cells include pThioHis A, B and C, pcDNA3.1/His, pEBVHis A, B and C, (Invitrogen, San Diego, Calif.), MPS V vectors, and numerous other vectors known in the art for expressing other peptides or proteins. Useful viral vectors include vectors based on retroviruses,
- 15 adenoviruses, adeno-associated viruses, herpes viruses, vectors based on SV40, papilloma virus, HBP Epstein Barr virus, vaccinia virus vectors and Semliki Forest virus (SFV). See, Brent et al., supra; Smith, 1995, Annu. Rev. Microbiol. 49:807; and Rosenfeld et al., 1992, Cell 68: 143. In particular, retroviral, lentiviral, adenoviral or adeno-associated viral vectors are commonly used for expression in immune cells such as T-cells. Examples of such
- 20 vectors include the SFG retroviral expression vector (see Riviere et al., 1995, Proc. Natl. Acad. Sci. (USA) 92:6733-6737).

In some embodiments, the vector is a retroviral or lentiviral vector. Optionally, the vector is an SFG retroviral vector. In some embodiments the vector is a lentiviral vector. Lentiviral vectors include self-inactivating lentiviral vectors (so-called SIN vectors).

- 25 The choice of expression vector depends on the intended host cells in which the vector is to be expressed. Expression vectors for mammalian host cells can include expression control sequences, such as an origin of replication, a promoter, and an enhancer (see, e.g., Queen, et al., 1986, Immunol. Rev. 89:49-68), and necessary processing information sites, such as ribosome binding sites, RNA splice sites, polyadenylation sites, and transcriptional
- 30 terminator sequences. These expression vectors usually contain promoters derived from mammalian genes or from mammalian viruses. Suitable promoters may be constitutive, cell type-specific, stage-specific, and/or modulatable or regulatable. Useful promoters include, but are not limited to, the metallothionein promoter, the constitutive adenovirus major late promoter, the dexamethasone-inducible MMTV promoter, the SV40 promoter, the MRP
- 35 polIII promoter, the constitutive MPS V promoter, the tetracycline-inducible CMV promoter (such as the human immediate-early CMV promoter), the constitutive CMV promoter, the

EF1 alpha promoter, the phosphoglycerate kinase (PGK) promoter and promoter-enhancer combinations known in the art.

5 Cultures of transformed organisms can be expanded under non-inducing conditions without biasing the population for coding sequences whose expression products are better tolerated by the host cells. In addition to promoters, other regulatory elements may also be required or desired for efficient expression of the amiRNA and CAR of the invention. These elements typically include an ATG initiation codon and adjacent ribosome binding site or other sequences. In addition, the efficiency of expression may be enhanced by the inclusion of enhancers appropriate to the cell system in use (see, e.g., Scharf et al., 1994, Results
10 Probl. Cell Differ. 20:125; and Bittner et al., 1987, Meth. Enzymol., 153:516). For example, the SV40 enhancer or CMV enhancer may be used to increase expression in mammalian host cells.

In some embodiments, a non-coding region of the vector comprises the polynucleotide. For example, a non-coding region of a retroviral vector may comprise the polynucleotide.

15 The genetic engineering of host cells can be carried out according to standard cloning and expression techniques, which are known in the art (e.g., as described in Sambrook, J., Fritsch, E. F., and Maniatis, T. Molecular Cloning: A Laboratory Manual 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989). The vector of the invention may be introduced into a host cell using such techniques.

20 The various forms of the term "transfection" are intended to encompass a wide variety of techniques commonly used for the introduction of exogenous DNA into a prokaryotic or eukaryotic host cell, e.g., electroporation, calcium-phosphate precipitation, DEAE-dextran transfection and the like.

Representative host cells include many *E. coli* strains, mammalian cell lines, such as CHO, CHO-K1, and HEK293; insect cells, such as Sf9 cells; and yeast cells, such as *S. cerevisiae*
25 and *P. pastoris*.

In some embodiments, the host cell is a mammalian host cell. For example, the host cell may be a mouse, rat, human, dog, cat, horse, cow, primate, goat or sheep host cell. In some embodiments, the host cell is a human or mouse host cell, preferably a human host
30 cell.

Cell lines which may be used include the NK cell line NK-92.

Mammalian host cells encompassed by the invention include Chinese Hamster Ovary (CHO cells) (including dhfr- CHO cells, described Urlaub and Chasin, 1980, Proc. Natl. Acad. Sci. USA 77:4216-4220 used with a DH FR selectable marker, e.g., as described in R.J. Kaufman

and P.A. Sharp, 1982, Mol. Biol. 159:601-621), NSO myeloma cells, COS cells and SP2 cells. In some embodiments the host cells are CHO K1PD cells. In some embodiments the host cells are NSO1 cells. In particular, for use with NSO myeloma cells, another expression system is the GS gene expression system shown in WO 87/04462, WO 89/01036 and EP 338,841.

In some embodiments, the host cell is an immuno-responsive cell. In the context of the present invention, an immuno-responsive cell is an immune cell. Typically, an immuno-responsive cell can be involved in an immune response, preferably an inflammatory immune response, for example to respond to cancer.

Also provided is a pharmaceutical composition comprising a peptide or a polynucleotide or a vector of the invention, and a pharmaceutically acceptable carrier.

The carrier and/or diluent is generally selected to be suitable for the intended mode of administration and can include agents for modifying, maintaining, or preserving, for example, the pH, osmolarity, viscosity, clarity, colour, isotonicity, odour, sterility, stability, rate of dissolution or release, adsorption, or penetration of the composition. Typically, these carriers and/or diluents include aqueous or alcoholic/aqueous solutions, emulsions, or suspensions, including saline and/or buffered media.

Suitable further agents for inclusion in the pharmaceutical compositions include, but are not limited to, amino acids (such as glycine, glutamine, asparagine, arginine, or lysine), antimicrobials, antioxidants (such as ascorbic acid, sodium sulphite, or sodium hydrogen-sulphite), buffers (such as borate, bicarbonate, Tris-HCl, citrates, phosphates, or other organic acids), bulking agents (such as mannitol or glycine), chelating agents (such as ethylenediamine tetraacetic acid (EDTA)), complexing agents (such as caffeine, polyvinylpyrrolidone, beta-cyclodextrin, or hydroxypropyl-beta-cyclodextrin), fillers, monosaccharides, disaccharides, and other carbohydrates (such as glucose, mannose, or dextrans), proteins (such as free serum albumin, gelatin, or immunoglobulins), colouring, flavouring and diluting agents, emulsifying agents, hydrophilic polymers (such as polyvinylpyrrolidone), low molecular weight polypeptides, salt-forming counterions (such as sodium), preservatives (such as benzalkonium chloride, benzoic acid, salicylic acid, thimerosal, phenethyl alcohol, methylparaben, propylparaben, chlorhexidine, sorbic acid, or hydrogen peroxide), solvents (such as glycerin, propylene glycol, or polyethylene glycol), sugar alcohols (such as mannitol or sorbitol), suspending agents, surfactants or wetting agents (such as pluronics; PEG; sorbitan esters; polysorbates such as Polysorbate 20 or Polysorbate 80; Triton; tromethamine; lecithin; cholesterol or tyloxapal), stability enhancing agents (such as sucrose or sorbitol), tonicity enhancing agents (such as alkali metal halides, such as sodium or potassium chloride, or mannitol sorbitol), delivery vehicles, excipients and/or pharmaceutical adjuvants.

The carrier and/or diluent may be a parenteral, optionally intravenous vehicle. Suitable parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride and lactated Ringer's. Suitable physiologically-acceptable thickeners such as carboxymethylcellulose, polyvinylpyrrolidone, gelatin and alginates may be included.

- 5 Intravenous vehicles include fluid and nutrient replenishers and electrolyte replenishers, such as those based on Ringer's dextrose. In some cases, one might include agents to adjust tonicity of the composition, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in a pharmaceutical composition. For example, in many cases it is desirable that the composition is substantially isotonic. Preservatives and other additives, 10 such as antimicrobials, antioxidants, chelating agents, and inert gases, may also be present. The precise formulation will depend on the route of administration. Additional relevant principle, methods and components for pharmaceutical formulations are well known (see, e.g., Allen, Loyd V. Ed, (2012) Remington's Pharmaceutical Sciences, 22nd Edition).

- A pharmaceutical composition of the present invention can be administered by one or more 15 routes of administration using one or more of a variety of methods known in the art. As will be appreciated by the skilled person, the route and/or mode of administration will vary depending upon the desired results. Routes of administration for pharmaceutical compositions of the invention include intravenous, intramuscular, intradermal, intraperitoneal, intrapleural, subcutaneous, intratumoural, spinal, or other parenteral routes 20 of administration, for example by injection or infusion. The phrase "parenteral administration" as used herein means modes of administration other than enteral and topical administration, usually by injection, and includes, without limitation, intravenous, intramuscular, intraarterial, intrathecal, intracapsular, intraorbital, intracardiac, intradermal, intraperitoneal, transtracheal, subcutaneous, subcuticular, intraarticular, subcapsular, 25 subarachnoid, intraspinal, epidural, intratumoural, intrapleural and intra-sternal injection and infusion. In some embodiments, the pharmaceutical composition is administered intratumourally. In other embodiments, administration is intrapleural or intraperitoneal. When parenteral administration is contemplated, the pharmaceutical compositions are usually in the form of a sterile, pyrogen-free, parenterally acceptable composition. A 30 particularly suitable vehicle for parenteral injection is a sterile, isotonic solution, properly preserved. The pharmaceutical composition can be in the form of a lyophilizate, such as a lyophilized cake.

- Alternatively, the pharmaceutical composition of the invention can be administered by a 35 non-parenteral route, such as a topical, epidermal, or mucosal route of administration, for example, intranasally, orally, vaginally, rectally, sublingually, or topically.

In some embodiments, the pharmaceutical composition is for subcutaneous administration. Typically, the pharmaceutical compositions for subcutaneous administration contain suitable

stabilizers (e.g., amino acids, such as methionine, and or saccharides such as sucrose), buffering agents and tonicifying agents. Alternatively, the pharmaceutical composition may be for intravenous administration.

5 Further provided is the peptide, polynucleotide, vector or pharmaceutical composition comprising the peptide of the invention, for use in therapy. In particular, the peptide, polynucleotide, vector or composition may be for use in the treatment of infections, especially bacterial infections, or the treatment of cancer.

10 The cancer may be any cancer such as, but not limited to cancers found in the GI tract, such as cancers of the GI tract, including colorectal carcinoma, cancers of the oral cavity; the brain, the lung, the kidney, the vagina, the bladder, the liver, the skin, the breast, the prostate and/or the ovary.

The infection may be any infection, particularly a bacterial infection, such as infections of the bladder, kidney or urinary tract.

15 Also provided is a method of treating an infection or cancer comprising the step of administering an effective amount of the peptide, polynucleotide, vector or pharmaceutical composition of the invention to a subject in need thereof.

Further provided is the peptide of the invention, or a pharmaceutical composition comprising the peptide of the invention, for use in the treatment of a disease modulated by L-MYC, N-MYC or C-MYC, or for use in regulating L-MYC, N-MYC or C-MYC.

20 Also provided is a method of treating a disease modulated by L-MYC, N-MYC or C-MYC comprising the step of administering an effective amount of the peptide or pharmaceutical composition of the invention to a subject in need thereof.

25 Further provided by the invention is a peptide comprising an amino acid sequence having substantial homology to the amino acid sequence of Lon Protease or a functional variant or fragment thereof for use in the treatment of a disease, such as an infection or cancer, wherein the disease is modulated by L-MYC or N-MYC.

Throughout the description and claims of this specification, the words "comprise" and "contain" and variations of the words, for example "comprising" and "comprises", mean "including but not limited to", and do not exclude other components, integers or steps.
30 Moreover the singular encompasses the plural unless the context otherwise requires: in particular, where the indefinite article is used, the specification is to be understood as contemplating plurality as well as singularity, unless the context requires otherwise.

In all instances, the terms "comprises" or "comprising" cover and can be replaced by "consists of" or "consisting of".

Preferred or optional features of each aspect of the invention may be as described in connection with any of the other aspects. Within the scope of this application, it is expressly intended that the various aspects, embodiments, examples and alternatives set out in the preceding paragraphs, in the claims and/or in the following description and drawings, and in particular the individual features thereof, may be taken independently or in any combination. That is, all embodiments and/or features of any embodiment can be combined in any way and/or combination, unless such features are incompatible.

"About" as used herein when referring to a measurable value such as an amount, a percentage, and the like, is meant to encompass variations of $\pm 20\%$ or $\pm 10\%$, more preferably $\pm 5\%$, even more preferably $\pm 1\%$, and still more preferably $\pm 0.1\%$ from the specified value, as such variations are appropriate to perform the disclosed methods. In all instances where the term "about" appears with reference to a measurable value, the disclosure includes the instance or embodiment without the term "about".

BRIEF DESCRIPTION OF THE DRAWINGS

One or more embodiments of the invention will now be described, by way of example only, with reference to the accompanying drawings, in which:

Fig. 1. c-MYC inhibition by rLon protease treatment.

a, Schematic of the protocol used for rLon protease treatment of *E. coli* CFT073 infected *Irf3*^{-/-}, *Irf7*^{-/-} and C57BL/6WT mice. Mice were treated with rLon, administered intraperitoneally (100 μ L, 250 μ g/mL) once a day for seven days starting one day before intravesical infection with *E. coli* CFT073. **b**, MYC-dependent genes regulated by infection in Placebo mice kidneys. 497 genes were activated, and 256 genes were inhibited. **c**, Volcano plot showing the strong activation of the total number of regulated genes in infected mice kidneys (red, upregulated; blue, downregulated; grey, not significantly regulated; cutoff $FC \geq 2$, $P < 0.05$ relative to uninfected mice). **d**, Inhibition of the response to infection by rLon treatment. Strong reduction in the number of MYC-dependent genes (2 genes were activated vs. 753 genes in Placebo), and the total number of regulated genes (43 regulated genes vs 3696 in Placebo), in the rLon treated mice. **e**, c-MYC protein staining in kidney tissues sections of rLon treated mice compared to the placebo group. Confocal microscopy; green, MYC; blue, nuclei; scale bar, 10 μ m). **f**, Quantification of c-MYC protein staining from **e** ($n = 3$ uninfected mice, $n = 7$ placebo 24h, $n = 5$ rLon 24h, $n = 3$ placebo 7d, $n = 4$ rLon 7d). Lines represent medians, One-way ANOVA followed by Dunn's correction. **g**, Inhibition of the innate immune response to infection by rLon treatment. Activation Z-score for each pathway is shown. **h**, Activation of MYC-dependent innate immune response in the placebo group. **i**, Inhibition of MYC-dependent innate immune response pathways and inflammation in rLon treated mice compared to the placebo group. Orange, upregulated; blue, downregulated.

Fig. 2. Therapeutic efficacy of rLon in a murine acute pyelonephritis model.

a, b, rLon protease treatment of *Irf3*^{-/-} mice (see schematic in Fig. 1) reduced gross kidney pathology compared to the placebo group (abscesses indicated by arrows, scale bars, 5mm, left kidneys). **c**, Improved bacterial clearance from renal tissues of rLon treated *Irf3*^{-/-} mice (right kidneys). **d**, Reduced bacterial counts in urine. **e**, Inhibited inflammation, quantified as neutrophil numbers in urine (red, placebo; blue, rLon). **f**, Results from the placebo group, showing the disease phenotype in infected *Irf3*^{-/-} mice, with massive neutrophil infiltration into infected kidneys and micro-abscess formation. Staining of whole kidney sections with hematoxylin and eosin (H&E). **g**, Significant increase in kidney size in infected *Irf3*^{-/-} mice compared to healthy kidneys is reversed in rLon treated mice. **h**, Tissue sections from rLon treated *Irf3*^{-/-} mice, showing smaller kidney size, pathology, and lower neutrophil infiltration, compared to the placebo group. **i**, Colocalization of bacteria and neutrophils in areas of pathology in the placebo group (*n* = 4 mice per group; red, neutrophil; green, bacteria; blue, nuclei; scale bar, see figure). **j**, Reduced bacterial and neutrophil staining in rLon treated mice, visualized by immunohistochemistry (*n* = 4 mice per group; red, neutrophil; green, bacteria; blue, nuclei; scale bar, see figure). **k**, Quantification of neutrophil and bacterial staining in kidney tissue sections from **i, j** Data are presented as mean ± s.d (**b, c, k**) or s.e.m (**d, e**), analyzed by Kruskal–Wallis with Dunn’s correction (**b, c, k**) or Two-way ANOVA followed by Bonferroni’s multiple comparison test (**d, e**). ****P* < 0.001.

Fig. 3. Inhibition of nephrotoxicity by rLon in *Irf3*^{-/-} mice and lack of toxicity in uninfected mice.

a, Expression of genes associated with renal toxicity, renal injury or kidney specific cell death in *E. coli* CFT073 infected *Irf3*^{-/-} mice. MYC- and IRF7-dependent genes regulating renal injury or kidney specific cell death were strongly upregulated in the placebo group compared to uninfected mice. **b**, Complete inhibition of genes regulating renal injury or kidney specific cell death in rLon treated mice compared to the placebo group. **c**, Inhibition of kidney damage, injury and cell death pathways by rLon treatment. **d**, Top regulated renal toxicity genes. Genes that were regulated in the placebo group were not regulated in the rLon treated group except *Lyz*. **e**, Principal component analysis of whole kidney mRNA profiles. Right cluster shows uninfected mice (green) and adjacent rLon treated mice (purple). Left cluster shows the placebo group (red). **f**, Heatmap of total detected genes and all renal toxicity genes, comparing uninfected controls to Placebo and rLon mice (fold change vs uninfected, white, no change; red, activated; blue, inhibited). Transcriptomic profile in the rLon treated group resembled the uninfected controls with pale colors. Placebo mice showed strong regulation of genes compared to uninfected mice.

Fig. 4. Involvement of L-Myc and N-Myc in the disease response to infection and cancer

- a-b**, Network analysis of L-Myc dependent genes regulated by *E. coli* CFT073 infection. rLon treatment inhibited the L-Myc response, with only one L-Myc dependent gene that was differentially expressed after 24 hours and none after seven days. **c**, Reduced L-Myc staining in kidney sections obtained from rLon treated mice, visualized by immunohistochemistry ($n = 4$ mice per group; green, L-Myc; blue, nuclei; scale bar, see figure). $n = 2$ mice per group, cutoff $FC \geq 2$, $P < 0.05$ relative to uninfected mice
- d-e**, Network analysis of N-Myc dependent genes regulated by *E. coli* CFT073 infection. rLon treatment inhibited the N-Myc response, with only 10 N-Myc dependent genes that were differentially expressed after 24 hours and 2 after seven days. **f**, Reduced N-Myc staining in kidney sections obtained from rLon treated mice, visualized by immunohistochemistry ($n = 4$ mice per group; green, L-Myc; blue, nuclei; scale bar, see figure). $n = 2$ mice per group, cutoff $FC \geq 2$, $P < 0.05$ relative to uninfected mice. **g**, Schematic of the protocol used for rLon protease treatment of colon cancer prone *Apc^{Min/+}* mice. Mice were treated with rLon, administered by oral gavage (100 μ L, 250 μ g/mL) twice daily for 14 days, starting at the age of 11 weeks. **h-i**, Strong activation of c-Myc, N-Myc and L-Myc dependent genes in *Apc^{Min/+}* mice compared to disease-free C57BL/6 controls. **k-m**, Reduced regulation and inhibition of c-Myc, N-Myc and L-Myc dependent genes in rLon-treated *Apc^{Min/+}* mice. $n = 2$ mice per group, cutoff $FC \geq 2$, $P < 0.05$ relative to healthy C57BL/6WT mice.

Fig. 5. LON-MYC molecular data

- a**, Structural overview of the Lon protease, highlighting the functional domains. **b**, Western blot of whole cell extracts of human kidney carcinoma cells, exposed to recombinantly expressed full-length Lon, the N-terminal domain (NTD), the ATPase domain (AAA) or the peptidase domain (Pep). Anti-Lon antibodies. **c**, Cellular uptake of AlexaFluor labeled full-length Lon, NTD, AAA or the peptidase domains in human lung carcinoma cells. Cytoplasmic uptake and perinuclear accumulation. Magenta, labelled proteins; blue, nucleus. **d-f**, Analysis of MYC levels in cells treated with full-length Lon, the N-terminal domain (NTD), the ATPase domain (AAA) or the protease domain (Pep). **d**, Reduced c-Myc levels in total cell extracts and nuclear extracts. **e**, Reduced N-Myc levels in total cell extracts and nuclear extracts. **f**, Reduced L-Myc levels in total cell extracts and nuclear extracts. **g**, PCR quantification of L-Myc, N-Myc and c-Myc mRNA in cells exposed to full-length Lon, NTD, AAA or the peptidase domain peptides. Pronounced inhibition of L-Myc (*MYCL* VA). Fold change compared to PBS control cells. **h-i**, *In silico* docking of the NTD peptide with c-Myc and N-Myc using Alfafold. Strong binding of NTD to the leucine-zipper rich domain of MYC is predicted (IPTM = 0.697).

Fig. 6. Cellular interactions of rLon and rLon domains and their effect on c-MYC, L-MYC and N-MYC expression

a, b, Interactions of the Lon NTD with the C-terminal part of c-MYC (residues 356-439,) and N-MYC (residues 354-464) as predicted by the AlphaFold multimer. **c,** Biacore analysis shows the dose-dependent interaction of full-length Lon and NTD peptide to His-tagged c-Myc and N-myc captured on Sensor Chip NTA and 1:1 Langmuir fitted curves. The fitted values (K_D) are indicated for each sensogram. **d,** Overlap of the NTD and MAX suggesting that the NTD targets the same binding site in MYC and MAX. **e,** Inhibition of c-MYC activation by the NTD and AAA+ as quantified by MYC/MAX complex binding to DNA in an ELISA-based assay. **f, g,** Strong inhibition of the generic transcription pathway by NTD and AAA+ in human kidney carcinoma cells treated with the NTD and AAA+ domain peptides (cutoff FC ≥ 1.5 , $n = 2$ experiments).

Fig. 7. Structure-function analysis of Lon protein interactions with MYC **a,** Structural overview of Lon highlighting the N-terminal domain, the central AAA+ domain and the C-terminal peptidase domain. **b,** Purity of the recombinant Lon, Nterminal domain (NTD), AAA+ and peptidase domains, verified by Western blot using anti-Lon antibodies. **c-e** Interactions of the Lon NTD (blue) with the C-terminal part of (c) c-MYC (residues 356-439, green), (d) N-MYC (residues 354-464, forest green). No predicted interaction of the Lon NTD with (e) L-MYC (lemon green) by the AlphaFold multimer. **f,** The SPR binding analysis shows the dose-dependent interaction of fulllength Lon and NTD peptide to His-tagged c-Myc and N-myc captured on Sensor Chip NTA (magenta) and 1:1 Langmuir fitted curves (black). The table shows the dissociation constant (K_D) of bindings. **g,** Overlap of the NTD and MAX suggesting that the NTD targets the same binding site in MYC and MAX. **h,** Inhibition of c-MYC activation by the NTD and AAA+ as quantified by MYC/MAX complex binding to DNA in an ELISA-based assay.

AlphaFold predicted interactions between full-length c-MYC and the full-length Lon protein showed a low iPTM score of 0.171. AlphaFold predicted interactions of fulllength c-MYC, N-MYC and L-MYC with Lon peptidase domain (aa 585-784). Putative complexes of all three MYC proteins and the peptidase domain resulted in relatively low iPTM scores (0.324, 0.32 and 0.3 for c-, N- and L-MYC, respectively).

Fig. 8. Cellular interactions of rLon and rLon domains and their effect on c-MYC, L-MYC and N-MYC expression

a, Live-cell confocal images showing the efficient uptake of AZ647-labeled (magenta) rLon (0.902 μ M) and rLon domains (6.314 μ M) into the cytoplasm, perinuclear area and nuclei of A498 cells. Nuclei were counterstained with Hoechst (blue) (Quantification of fluorescence intensity shown in each image; 50 cells per sample; $n = 2$ experiments; scale bars, 3 μ m).

b, Western blot confirming the uptake of rLon and rLon domains in whole-cell extracts from A498 human kidney carcinoma cells exposed to rLon and its domain peptides. **c-e**, Confocal images showing inhibition of c-MYC and N-MYC protein levels in Lon and NTD-treated cells compared to control. L-MYC protein level was also reduced in peptidase-treated cells compared to control. (green, c-MYC, N-MYC or L-MYC; blue, nuclei; scale bars, 10 μ m). Lower panel: quantification of fluorescence intensity (50 cells per sample; $n = 2$ experiments). Data are presented as the mean $\pm$ s.d. **f**, Strong inhibition of *MYCL* mRNA expression in NTD-treated A498 cells. *MYCN* expression was weakly inhibited by the full-length Lon protein and the NTD peptide. *MYC* expression was not affected. *Irf7* was used as a negative control. Data are presented as the mean $\pm$ s.d. ($n = 2-3$ experiments for *MYC*, *MYCL* and *MYCN*; $n = 1$ experiment for *IRF*). **g, h**, Strong inhibition of the generic transcription pathway by NTD and AAA+ in human kidney carcinoma cells treated with the NTD and AAA+ domain peptides (cutoff FC ≥ 1.5 ; $n = 2$ experiments).

Fig. 9. MYC degradation by Lon and Lon peptide domains a-c, Analysis of protein degradation by rLon (a) and the peptidase and AAA+ domains (b). c-MYC (50 ng) was degraded by rLon (0.5 or 1 μ g) after 120 min and by the rLon peptidase domain, in the presence of the AAA+ domain peptide. The c-MYC protein control is shown with and without incubation. This effect was reversed in the presence of protease inhibitors (c). **d, e**, Inhibition by rLon (d) and the NTD peptide (e) of c-MYC, N-MYC and L-MYC protein expression in whole-cell and nuclear extracts from A498 cells. **f**, Quantification of Western blot as percent of inhibition compared to untreated cells. Data are presented as the mean. ($n = 1-3$ experiments).

Fig. 10. Rapid c-MYC response to kidney infection and inhibition by rLon protease treatment.

a, Schematic of the protocol used for rLon treatment of *E. coli* CFT073-infected *Irf3*^{-/-} mice. Mice were administered rLon intraperitoneally (100 μ L, 250 μ g/ml) once a day for eight days starting one day before intravesical infection with *E. coli* CFT073. Total renal RNA was isolated from infected mice or uninfected mice after 24 hours and seven days and was subjected to genome-wide transcriptomic analysis and disease-associated genes were identified by comparing the infected mice to uninfected controls. **b, c**, Response to kidney infection in the placebo group, showing activation of MYC-dependent genes after 24 hours and seven days ($n = 2$ per group; red, upregulated; blue, downregulated; cutoff FC ≥ 2 , $P < 0.05$ compared to uninfected mice). **d**, Inhibition by rLon treatment of the MYC-response to infection, shown as a marked reduction in the number of MYC-dependent genes in rLon-treated mice compared to the placebo group. **e**, c-Myc protein staining in kidney tissue sections of the rLon-treated group compared to the placebo group. Confocal microscopy; representative images; green, c-Myc; blue, nuclei; scale bar, 20 μ m; $n = 3-7$ mice per

group. **f**, Quantification of c-Myc protein staining from **e**. Lines represent medians. One-way ANOVA followed by Dunn's correction. **g, h**, Functional analysis of MYC dependent genes, identifying the innate immune response as activated in the placebo group and inhibited in the rLon-treated group. Orange, upregulated; blue, downregulated. **i**, Inhibition of the innate immune response to infection by rLon treatment. The activation Z score for each pathway is shown. Data normalized against healthy, uninfected mice.

Fig. 11. Activation of L-Myc and N-Myc is inhibited by rLon treatment in infection and cancer models **a, b**, Network analysis of L-MYC-dependent genes regulated by *E. coli* CFT073 infection in *Irf3*^{-/-} mice. The L-Myc response to kidney infection was inhibited by rLon treatment, with only one L-MYC-dependent gene differentially expressed after 24 hours and none after seven days ($n = 2$ mice per group; cutoff $FC \geq 2$; $P < 0.05$ compared to uninfected mice). **c, d**, The N-Myc response to kidney infection was inhibited by rLon treatment, with only ten N-MYC-dependent genes differentially expressed after 24 hours and two after seven days ($n = 2$ mice per group; cutoff $FC \geq 2$; $P < 0.05$ compared to uninfected mice). **e, f**, Reduced L-Myc staining and N-Myc staining in kidney sections obtained from rLon-treated mice visualized by immunohistochemistry (Representative images; $n = 2-5$ mice per group; green, N-Myc or L-Myc; blue, nuclei). **g**, Schematic of the protocol used for rLon treatment of colon cancer-prone *Apc*^{Min/+} mice by oral gavage (100 μ l, 250 μ g/ml) twice daily for 14 days, starting at 11 weeks of age. **h-j**, Strong activation of c-MYC-, N-MYC- and L-MYCdependent genes in *Apc*^{Min/+} mice compared to disease-free C57BL/6 control mice. **km**, Inhibition of c-MYC-, N-MYC- and L-MYC-dependent genes in rLon-treated *Apc*^{Min/+} mice ($n = 2$ mice per group; cutoff $FC \geq 2$; $P < 0.05$ relative to disease-free C57BL/6 control mice).

Fig. 12. Inhibition of nephrotoxicity by rLon in *Irf3*^{-/-} mice and proximity of treated mice to uninfected mice.

a, Gene expression analysis identifying genes associated with renal toxicity, renal injury or kidney-specific cell death in *E. coli* CFT073-infected *Irf3*^{-/-} mice. Strong activation of MYC- and IRF7-dependent genes regulating renal injury or kidney-specific cell death in the placebo group compared to the uninfected group. Red, upregulated, blue, downregulated; cutoff $FC \geq 2$; $P < 0.05$ compared to the uninfected group. **b**, Complete inhibition of genes regulating renal injury or kidney-specific cell death in the rLon-treated group compared to the placebo group. **c**, Inhibition of kidney damage, injury and cell death pathways by rLon treatment compared to placebo. **d**, Top regulated renal toxicity genes in the placebo group were not regulated in the rLontreated group, with the exception of *Lyz*. **e**, Principal component analysis of the whole kidney RNA data sets. The right cluster shows uninfected mice (green) and the adjacent cluster shows the rLon-treated mice (purple). The left cluster shows the placebo group (red). **f-h**, Volcano plots of all the detected genes (**f**), all

MYC-dependent genes (**g**) and all renal toxicity genes (**h**) ($n = 2$ per group; FC vs. uninfected; white, no change; red, activated; blue, inhibited). Two gene categories were distinguished in the placebo group; category one was the strong disease response ($n = 3,695$) and category two was the remaining genes that were not regulated by disease ($n = 16,168$). Category one (disease related genes) was inhibited in rLon-treated mice ($n = 43$) but category two (genes unrelated to disease) was not significantly regulated by rLon treatment ($n = 19,820$). The MYC-dependent genes and renal toxicity genes showed a similar profile in the rLon-treated group and resembled that in the uninfected control group.

10 **Figure 13. Supplementary Data for Fig. 1**

Irf3^{-/-} mice were infected with the uropathogenic *E. coli* strain CFT073, and kidney tissues were obtained for gene expression analysis after 24 hours and seven days. **a- c**, Heatmap (**a**), histogram (**b**) and Venn diagram (**c**) comparing the gene expression profiles of infected mice after 24 hours and seven days to those of uninfected mice ($n = 2$ per group; red, upregulated; blue, downregulated; cutoff $FC \geq 2$, $P < 0.05$ compared to uninfected mice). **d**, Volcano plots of whole gene expression in infected *Irf3*^{-/-} mice after 24 hours and a further increase in gene expression after seven days.

2

e, Activation of *Myc* and MYC-dependent gene networks in infected *Irf3*^{-/-} mice after 24 hours, including an innate immune response network. Inhibition of the innate immune response to infection by rLon treatment (orange, activation; blue, inhibition). Data normalized against healthy, uninfected *Irf3*^{-/-} mice.

20 **Figure 14. Supplementary Data for Figure 1**

c-Myc protein staining in kidney tissue sections of the rLon-treated group compared to the placebo group. Fluorescence microscopy; representative images shown in Fig. 1; green, c-Myc; blue, nuclei; scale bar, 20 μm ; $n = 3-4$ mice per group.

Figure 15. Supplementary Data for Figure 1.

Pathway analysis showed that interferon signaling, Toll-like receptor signaling, NF- κ B signaling and inflammasome pathways were inhibited by rLon treatment in *Irf3*^{-/-} mice. ($n = 2$ per group; red, upregulated; blue, downregulated; cutoff $FC \geq 2$, $P < 0.05$ compared to uninfected mice).

Figure 16. Supplementary Data for Figure. 2.

a, f, k, Macroscopic appearance of kidneys from *Irf3*^{-/-} mice in the placebo or rLon treatment group. **b-e, g-j, l-o**, Lower gross pathology, urine neutrophil count, and bacterial count in urine and renal tissues were compared in mice treated with rLon or placebo. The results of three separate experiments are shown. *n* = 5 mice per treatment for the 1st and 2nd experiments; *n* = 3 mice per treatment for the 3rd experiment. Lines represent medians (**b, c, g, h, l, m**), and data were analyzed by the Kruskal-Wallis test with Dunn's correction. Data are presented as the mean ± s.e.m. (**d, e, i, j, n, o**) after analysis by two-way ANOVA with Bonferroni's multiple comparison test. **P* < 0.05, ***P* < 0.01, ****P* < 0.001, n.s. = not significant.

10 **Figure 17. Supplementary Data for Figure 2.**

Staining of whole kidney sections with hematoxylin and eosin (H&E) showing the massive neutrophil infiltration into infected kidneys and microabscess formation and no pathological characteristics and lower neutrophil infiltration in rLon-treated mice (Representative images shown in Fig. 2; *n* = 4 mice per group; scale bar, 500 μm.).

15 **Figure 18. Control experiments for Figure 2.**

a, b, The Lon effect on growth of *E. coli* CFT073 *in vitro*. *E. coli* CFT073 were cultured in Luria Broth media (black) for 3.5 hours followed by addition of 100 mg/ml of Cefotaxime (red) or 250 μg/ml (dark blue) or 500 μg/ml (blue) of rLon. OD600 was measured at 0, 1, 2, 3, 4 and 5 hours. rLon did not affect the growth of *E. coli* CFT073. Two repeat experiments for each rLon concentration used together with Cefotaxime as the positive control. **c, d**, Cell death assays for A498 kidney carcinoma cells based on Presto Blue (**c**) and ATP Lite (**d**) assays for full-length rLon and NTD, AAA+ and peptidase domain peptides. No toxicity was detected for full-length rLon, NTD, AAA+ nor peptidase domain peptide.

Figure 19. Supplementary Data for Figure. 9.

25 **a**, High-resolution confocal imaging of renal tissue sections from infected *Irf3*^{-/-} mice showing intense staining for bacteria and neutrophils and colocalization in areas of pathology (*n* = 4 mice per group; red, neutrophils; green, bacteria; blue, nuclei). **b**, Split channel images from **a**.

Figure 20. Supplementary Data for Figure 4.

N-Myc protein staining in kidney tissue sections of the rLon-treated group compared to the placebo group. Confocal microscopy; representative images shown in Fig. 4; green, N-Myc; blue, nuclei; scale bar, 20 μ m; n = 2-5 mice per group.

Figure 21. Supplementary Data for Figure 4.

- 5 L-Myc protein staining in kidney tissue sections of the rLon-treated group compared to the placebo group. Confocal microscopy; representative images shown in Fig. 4; green, L-Myc; blue, nuclei; scale bar, 20 μ m; n = 2-5 mice per group.

Figure 22. Supplementary Data for Figure 5.

- a-c**, Analysis of N-Myc protein degradation by rLon and the peptidase and AAA+ domains. N-MYC (50 ng) was degraded by rLon (0.5 or 1 μ g, **a**) after 120 min and by the rLon peptidase domain, in the presence of the AAA+ domain peptide (**b**). The N-MYC protein control is shown with and without incubation. This effect was reversed in the presence of protease inhibitors (**b**). **d**, Purity of the recombinant Lon, N-terminal domain (NTD), AAA+ and peptidase domains, verified by Coomassie staining. **e-g**, Western blot analysis showing the uptake of Lon and different domain peptides in different cellular fractions.

Figure 23. Supplementary Data for Figure 5.

- a**, Western blot analysis of c-MYC degradation by Lon in whole-cell lysates and nuclear extracts. **b**, Western blot analysis of N-MYC degradation by rLon in whole-cell lysates and nuclear extracts. **c**, Western blot analysis of L-MYC degradation by rLon and different domains in whole-cell lysates and nuclear extracts.

Figure 24. Supplementary Data for Figure 6.

- Biacore data the dose-dependent interaction of full-length Lon and NTD peptide to His-tagged c-Myc and N-myc captured on Sensor Chip NTA. The cell without nickel loading served as a reference for nonspecific binding of proteins and chemicals (blue), and buffer only (green) served as control. The Histidine-tagged ligand (c-Myc or N-myc) was injected over the nickel activated sensor surface to be captured. Full-length Lon or NTD peptide was then injected over the captured His-tagged ligand with five different dilutions for monitoring the interaction of sample with His-tagged ligand (magenta).

DETAILED DESCRIPTION

EXAMPLES**Example 1****Methods****5 Bacterial strains**

The prototype APN strains *E. coli* CFT073 (O6:K2:H1) (ref) were cultured on tryptic soy agar (TSA) plates (16 hours, 37°C), harvested in phosphate-buffered saline (PBS, pH 7.2, 10¹⁰ CFU/mL) and diluted as appropriate. Overnight static cultures of *E. coli* CFT073 in Luria broth were used for animal inoculations.

10 Experimental urinary tract infection

Mice were bred at the MIG animal facility, Lund, Sweden and housed in specific pathogen-free individually ventilated cages (IVC) at a constant temperature of 23°C on a 12-h light-dark cycle with lights on at 7:00 a.m, with food and water *ad libitum*. Female C57BL/6 wild type, *Irf3*^{-/-} and *Irf7*^{-/-} and mice were used at 9-15 weeks of age. *Irf7*^{-/-} and *Irf3*^{-/-} mice on a
15 C57BL/6 background were kindly provided by the Riken Bioresource Center, Japan, with permission from T. Taniguchi.

Mice anesthetized with Isoflurane were infected by intra-vesical inoculation with *E. coli* CFT073 (10⁸ CFU in 0.1 mL), and sacrificed under anesthesia. Kidneys were aseptically removed, and macroscopic pathology was documented by photography. Scoring was based
20 on the degree of hyperaemia, oedema and abscesses on a scale from 0-10, where 0 was similar to that of uninfected controls, and 10 had the most severe hyperemia, oedema and abscesses. One kidney was divided for RNA extraction (on dry ice) and immunostaining (in O.C.T. compound or 10% (v/v) formalin). Viable counts were determined by growth on TSA
25 plates (37°C, overnight) after serial dilutions as appropriate. Neutrophils in uncentrifuged urine were counted using a hemocytometer.

Colon cancer model

C57BL/6J-*Apc*^{Min}/J mice developing spontaneous intestinal adenoma formation were bred at the Department of Laboratory medicine and used at the age of either 11 or 20 weeks. The mice were treated twice daily for 14 days by oral gavage (200 µl) of rLon protease (250 µ
30 g ml⁻¹). At sacrifice, intestines were aseptically harvested and a segment of the intestine were then prepared for mRNA extraction for further analysis.

Bacterial Lon protease treatment *in vivo*

C57BL/6 wild-type, *Irf3*^{-/-}, and *Irf7*^{-/-} mice were treated with the recombinant bacterial Lon protease, administered intraperitoneally (100 µL, 250 µg/mL) once a day for 7 days starting one day before intravesical infection with *E. coli* 073 at day 0. The mice were sacrificed after
5 24 hours and 7 days. Sham treated mice received PBS.

Whole-genome transcriptomics analysis

Total renal RNA was isolated from using the RNAeasy Mini Kit (Qiagen), amplified using GeneChip 3'IVT Express Kit (Affymetrix), hybridized onto Mouse Genome 430 PM array strips and scanned in-house using the GeneAtlas System (Affymetrix). Data was normalized
10 using Robust Multi Average implemented in the Transcriptome Analysis Console software (v.4.0.1.36, Applied Biosystems). Significantly altered genes were calculated by comparing cells or tissue infected with bacteria to uninfected controls (PBS control) and sorted by relative expression (2-way ANOVA model using Method of Moments, absolute fold change > 1.41, 1.5 or 2.0). Heat maps were constructed using GraphPad (Prism v. 9). Differentially
15 expressed genes were functionally characterized using ingenuity pathway analysis software (IPA, Qiagen).

Immunohistochemistry

Tissues embedded and frozen in O.C.T. were cryosectioned (8 µm, Leica microtome), collected on positively charged microscope slides (Superfrost/Plus; Thermo Fisher
20 Scientific), fixed in acetone-methanol (1:1, 10 min), dried, permeabilized (0.2% Triton X-100, 5% normal goat serum/PBS) and stained with primary rat anti-neutrophil-antibody [NIMP-R14] (1:200; Abcam, ab2557), rabbit monoclonal anti-c-Myc antibody (1:100; ab32072, Abcam), L-Myc and N-Myc antibodies followed by Alexa 488 or Alexa 568 labeled rabbit-anti-rat and goat-anti-rabbit IgG secondary antibodies (Molecular Probes, A-21210,
25 A-11001 and A-11011). Nuclei were counterstained with DAPI (0.05 mM; Sigma-Aldrich). Slides were examined by fluorescence microscopy (AX60, Olympus Optical). Richard-Allan Scientific Signature Series Hematoxylin 7211 and Eosin-Y 7111 (Thermo Scientific) were used to counterstain the tissue sections.

Human kidney epithelial cells

30 The A498 human kidney epithelial cells (A498, ATCC HTB-44) were cultured in RPMI-1640 supplemented with 1 mM sodium pyruvate, 1 mM non-essential amino acids and 10% fetal bovine serum (FBS) overnight at 37°C, 90 % humidity and 5 % CO₂ in 6-well plates (for western blots), or 8-well chamber slides (for confocal imaging). Cells were then infected with bacteria in fresh, serum-free supplemented RPMI.

siRNA transfection

A498 human kidney epithelial cells were siRNA transfected, using Lipofectamine® RNAiMAX reagent, according to the manufacturer's instructions. Knockdown efficiency was validated by qRT-PCR. After a 48-h incubation, transfected cells were harvested for protein
5 extractions or confocal microscopy.

Immunofluorescence

A498 cells were fixed (3.7% formaldehyde, 10 min), permeabilized (0.25% Triton X-100, 5% FBS, 15 min), blocked (5% FBS, 1h at RT), incubated with anti-c-Myc (ab32072, Abcam), anti-IRF7 (Abcam), anti-Phospho-IRF3 or anti-Total IRF3 primary antibody (1:200)
10 in 5% FBS overnight at 4°C and Alexa Fluor® 488 goat anti-rabbit IgG secondary antibody (1:200; A-11034, Life Technologies) for 1 h at RT. After nuclear staining (DRAQ5, Abcam), slides were mounted (Fluoromount, Sigma-Aldrich), imaged by laser-scanning confocal microscopy (LSM510 META confocal microscope, Carl Zeiss) and quantified by ImageJ software 1.46r (NIH).

Western blotting

Cells were cultured in 6 well plates (4×10^5 cells/well, Thermo Fisher Scientific), washed with ice-cold PBS and lysed with RIPA buffer containing protease and phosphatase inhibitors (Roche Diagnostics). Cell lysates were run on SDS-PAGE (4-12% Bis-Tris gels, Invitrogen) and blotted onto PVDF membrane (GE Healthcare), blocked with 5% bovine serine albumin
20 (BSA) and incubated with anti-c-Myc (ab32072, Abcam), anti-IRF7 (Abcam), anti-Phospho-IRF3 or anti-Total IRF3 primary antibody (1:1000) primary antibody (1:1000 in 5% BSA). The blots were washed with PBS tween 0.1% (PBST) and incubated with secondary antibodies in 5% non-fat dry milk or BSA (1:4000, goat anti rabbit-HRP, Cell Signaling), then washed again with PBST, developed with ECL Plus detection reagent (GE Health Care)
25 and imaged using the Bio-RAD ChemiDoc™ system. β -actin (1:4,000, A5441, Sigma) was used as loading control.

Quantification and Statistical Analysis

For data determined to follow a Gaussian distribution (D'Agostino & Pearson normality test), statistical differences were determined by two-tailed *t*-test or one-way ANOVA with
30 Greenhouse–Geisser correction. For non-parametric data, either the two-tailed Mann–Whitney U-test or Kruskal–Wallis test with Dunn's correction was used and paired data were analyzed by Wilcoxon matched-pairs signed-rank test. Statistical significance was determined by GraphPad (Prism v. 9) and significance was assigned at * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$. All data are presented as the mean, and all error bars represent

s.d. Sample size (*n*) indicates either the number of independent experiments or the number of mice used in the study.

Results

A rapid MYC response to infection in disease-prone *Irf3*^{-/-} mice

5 The MYC transcription factors define essential aspects of renal growth and development and in childhood, acute pyelonephritis has been associated with renal scar formation, impaired renal growth and lasting detrimental effects on renal function. Female *Irf3*^{-/-} mice were infected with the uropathogenic *E. coli* strain CFT073 (Fig. 1). Mice infected by intra-vesical injection, developed severe kidney disease with elevated gross pathology scores after seven
10 days, characteristic of acute pyelonephritis. Bacterial counts remained high after seven days and by histopathology, renal tissues showed evidence of neutrophil infiltration and abscess formation, compared to uninfected mice (Fig. 2).

A rapid C-MYC response to infection was detected by gene expression analysis of renal tissue RNA. Total gene expression was elevated in infected mice after 24 hours (1093
15 significantly regulated genes) as well as MYC dependent gene networks and gene expression remained activated after seven days, including MYC and MYC-dependent genes (cut off FC > 2, *P* < 0.05 compared to uninfected controls). A parallel increase in renal C-Myc protein levels was detected by immunohistochemistry of tissue sections from infected *Irf3*^{-/-} mice, compared to uninfected controls. C-MYC staining was higher along the mucosal
20 lining of the renal pelvis and in the renal papilli, consistent with a local MYC response at the site of infection (Fig. 2). C-Myc-dependent regulators of innate immunity were strongly upregulated in infected *Irf3*^{-/-} mice compared to healthy controls (Fig. 1g), including toll-like receptors NFkB, IL-6 and type I interferons, as well as IL-8 and S100A8, which activate neutrophil recruitment.

25 The results identify a strong disease-associated C-Myc response in the kidneys of infected *Irf3*^{-/-} mice.

rLON treatment inhibits MYC overexpression in infected mice.

The inventors subsequently investigated if rLon acts as a MYC inhibitor in infected *Irf3*^{-/-} mice and if rLon treatment may have a protective potential. *Irf3*^{-/-} mice were intra-
30 peritoneally injected with recombinant Lon protease (rLon, 250 µg ml⁻¹), one day before infection and daily for seven days (Fig. 1a). The placebo group received PBS. rLON treatment reduced the total number of regulated genes from 3696 in the placebo group to 43 in the rLon treated group (day 7, cut off FC ≥ 2, *P* < 0.05) (Fig. 1d). C-Myc-overexpression was virtually abolished, as shown by a decrease from 747 deregulated MYC
35 genes in the placebo group to two genes in the rLon treated group (day 7, cut off FC ≥ 2, *P*

< 0.05, Fig. 1b, d). A parallel decrease in C-MYC staining was detected by immunohistochemistry of renal tissue sections from rLon treated, infected *Irf3*^{-/-} mice, compared to the placebo group, confirming the treatment effect at the protein level (Fig. 1e, f).

- 5 The innate immune response to infection was markedly reduced by rLon treatment, including a MYC dependent network of innate immune response regulators, which was strongly inhibited in rLon treated *Irf3*^{-/-} mice and major innate immune signaling pathways, activated in the placebo group, were no longer regulated in the treated mice, including toll-like receptor-, IL-6-, IL-8-, TNFR, NF-κB and inflammasome signaling (cut off FC ≥ 2, P < 10 0.05, Fig. 1 hIrf7-dependent genes were inhibited and interferon signaling was strongly attenuated.

rLon treated mice are protected from severe acute kidney disease

- rLon treatment was further shown to prevent infected *Irf3*^{-/-} mice from developing acute renal disease (Fig. 2). rLon treatment inhibited the development of gross pathology (Fig. 2a, 15 b) and bacterial numbers were reduced in urine and infected kidneys, compared to the placebo group (Fig. 2c, d). Neutrophil infiltration and abscesses formation was inhibited and neutrophil counts in urine were reduced (Fig. 2e). After seven days the renal tissue structure remained largely intact in the rLon treated group, in contrast to the placebo group (Fig. 2f-h).
- 20 This difference between the rLon treated mice and the placebo group was further investigated by high-resolution confocal imaging of tissue sections from infected rLon treated mice, compared to the placebo group (Fig. 2i-k). Sites of infection were identified by prominent co-localization of bacteria and neutrophils, indicating abscess formation. Intense staining for bacteria and neutrophils was detected in the mucosal lining of the renal pelvis, 25 in the papilli and collecting ducts and in foci adjacent to the apex of the renal pelvis (Fig. 2i). This staining pattern was not observed in the rLon treated mice, where staining for bacteria and neutrophils was low or not detected (Fig. 2j, k).

The results suggest that rLon treatment inhibits MYC dysregulation by infection and prevents the development of acute renal disease in susceptible mice.

30 rLon treated mice are protected from kidney damage.

- Gene expression analysis was used to examine the renal tissue damage caused by infection, in more detail (Fig. 3). A significant number of renal damage-associated genes were activated in the placebo group, compared to uninfected mice, consistent with the extensive renal pathology, observed in these mice (Fig. 3a). In contrast, the renal damage response 35 was virtually absent in the rLon treated group, with only one significantly regulated gene

associated with renal toxicity ($n = 1$) and no evidence of a broader MYC-dependent response (Fig. 3b). A large number of renal damage associated genes in the placebo group were definable as MYC-related (34%), including a subset of IRF7-dependent genes (NF- κ B, STAT3, FOS and IRFs). A further subset of renal damage associated, IRF7-related but MYC independent genes included TLRs, MYD88 and CD40 (Fig. 3a).

The results suggest that MYC dysregulation drives renal toxicity in infected mice, which includes genes involved in renal toxicity, renal injury or kidney specific cell death. rLon treatment prevented renal damage, by targeting and inhibiting this response.

rLon treatment targets deregulated MYC

The potent protective effect of rLon treatment in infected kidneys raised the question if rLon mainly targeted the deregulation of MYC that accompanies disease in infected kidneys. Principal Component Analysis of renal tissue RNA was used to evaluate the entire RNA data sets in the treatment and placebo groups, compared to healthy, uninfected mice. In contrast to the placebo group, which formed a cluster at a significant distance from the healthy mice, the rLon treated group was shown to closely resemble uninfected mice (Fig. 3f).

The effect of rLon on the MYC deregulation in infected kidneys was further analyzed by a direct comparison of the placebo group and rLon treated group to healthy, uninfected mice (Fig. 3e). rLon treatment was shown to specifically inhibit the disease- associated response, which characterized the placebo group, while retaining the homeostatic MYC expression, to a similar extent as healthy uninfected mice.

A similar analysis was performed for the subsets of Myc dependent genes in these data sets. The results show that rLon treatment mainly targets the deregulation of MYC that accompanies disease in infected kidneys.

L-MYC and N-MYC responses to infection are inhibited by rLon treatment

The C-MYC, L-MYC and N-MYC family members are functionally similar but show distinct expression patterns and tissue specificities. L-MYC has been localized to the renal pelvic mucosa and the papilli, where infection was localized in this study. N-MYC is also expressed during renal development. The inventors therefore examined if the kidney responses to infection also includes N-MYC and L-MYC and if rLon treatment affects L-MYC and N-MYC levels.

N-Myc and L-Myc expression was strongly activated in infected *Irf3*^{-/-} mice, after 24 hours and seven days, compared to uninfected controls. In the placebo group, activated N-Myc gene networks included about 25% of all genes defined by IPA as N-MYC-related (Fig. 4). A

similar activation pattern was detected for L-MYC, including about 33% of genes definable as L-MYC related. The increase in N-MYC and L-MYC protein levels was confirmed by immunohistochemistry in renal tissue sections, where increased L-Myc staining was detected along the renal pelvis and in the papilli after 24 hours and after seven days infected *Irf3*^{-/-} mice (Fig. 4c, f).

rLon treatment was further shown to inhibit the L-Myc and N-Myc response to infection in *Irf3*^{-/-} mice. L-Myc expression was reduced by 95% compared to the placebo group (number of genes) (Fig. 4b, e). Only one L-MYC dependent gene was differentially expressed after 24 hours and none remained activated after seven days (cut off FC>2, p<0.05). Only two N-MYC-related genes were differentially expressed after seven days of rLon treatment, compared to 400 genes in the placebo group. Treatment was further accompanied by a loss of L-MYC- and N-MYC staining in tissues from rLon treated mice (Fig. c, f).

The results suggest that N-Myc and L-Myc are deregulated in the kidneys of infected *Irf3*^{-/-} mice and that in addition to targeting C-Myc, rLon treatment targets N-MYC and L-Myc responses, which are inhibited.

rLon treatment inhibits N-Myc and L-Myc expression in a model of intestinal cancer

Peroral rLon treatment was previously shown to inhibit c-MYC expression in colon cancer prone *Apc*^{Min/+} mice and to prevent tumor progression and mortality. Here, intestinal sections were examined to investigate if N-Myc and L-Myc expression is increased in mice with intestinal tumors (Fig. 4h, m). Gene expression analysis detected significant activation of N-Myc and L-Myc expression in tumor biopsies from the placebo group and N-MYC and L-MYC dependent gene expression networks were activated in tumor bearing *Apc*^{Min/+} mice, compared to healthy C57BL/6 controls.

In contrast, the expression of N-Myc and L-Myc was not activated in intestinal tissue sections from rLon treated *Apc*^{Min/+} mice, compared to healthy C57BL/6 controls. Gene expression was significantly reduced in the rLon treated *Apc*^{Min/+} mice, compared to the placebo group.

The results identify rLon as a pan-MYC inhibitor in models of bacterial infection and cancer.

Cellular interactions of rLon and rLon peptide domains

The Lon protease is responsible for protein quality control in a variety of organisms including bacteria, where it hydrolyzes unfolded and misfolded proteins. Furthermore, Lon has been shown to interact with DNA. The effects of Lon in the kidney infection model supported the relevance of these two mechanisms, as MYC protein levels were reduced and the expression of MYC dependent genes strongly inhibited in treated mice. In analogy with

its effects on non-mammalian proteins, the Lon protease domain might be responsible for MYC protein degradation and a second domain for transcriptional inhibition.

To execute these effector functions, the Lon protein enters human cells and interacts directly through its domains with the MYC protein isoforms or the transcriptional machinery.

- 5 To identify Lon protein domains involved in MYC inhibition, globular Lon peptide domains were recombinantly expressed and human kidney carcinoma cells were exposed to the full-length rLon protein, the N-terminal domain peptide (NTD, residues 1-245), the ATPase-domain peptide (AAA, residues 246-584) or the peptidase domain peptide (residues 585-784) (Fig. 5a). The Alexa-Fluor labeled full-length protein and the peptides were efficiently
10 internalized as shown by live-cell confocal imaging and were detected in the cytoplasm, perinuclear- and nuclear areas of the cells (Fig. 5c). The uptake of Lon and the peptides was confirmed by Western blot analysis of whole cell extracts (Fig. 5b). The peptides were further localized to different compartments, by sub-cellular fractionation.

- A reduction of L-MYC, N-MYC and C-MYC was demonstrated by analysis of subcellular
15 fractions of cells treated with the full-length protein or each of the peptides. The reduction in MYC levels mainly affected nuclear protein levels (Fig. 5d, f). The full-length Lon affected all three MYCs, the N-terminal domain L-MYC, the AAA domain L-MYC, the peptidase domain N-MYC and L-MYC.

- Effects of Lon and the peptide domains on MYC gene expression were further investigated
20 by qRT-PCR, using primers specific for C-MYC, L-MYC (isoforms I and II) or N-MYC (Fig. 5g). L-MYC expression was markedly reduced in human kidney carcinoma cells exposed to the N-terminal domain peptide or the full-length Lon protein. A significant but less potent effect of N-terminal domain (NTD) was observed for N-MYC. C-MYC, however, was not significantly affected. The peptidase domain showed a significant inhibitory effect on L-MYC
25 expression suggesting that L-MYC is targeted transcriptionally to a greater extent than C-MYC.

MYC domains targeted by Lon and the domain peptides were further studied in silico

- AlfaFold was used to predict interactions with C-MYC, N-MYC and L-MYC (Fig. 5h, i).
30 Successful in silico docking predicted that the C-terminal part of C-MYC (aa 356-439) interacts with the LON NTD, with an iptm score of 0.697. Furthermore, there was a clear indication that the C-terminal part of N-MYC (aa 354-464), interacts with Lon with a similar binding mode and iptm score. The targeted MYC sequences corresponded to the alpha-helical leucine-rich zipper domains that bind DNA and regulate transcription, especially the
35 expression of MYC-dependent genes. L-MYC was not targeted by the NTD.

The results suggest that the NTD regulates cellular MYC levels, in part by binding to alpha-helical leucine-rich zipper domains.

Disease specificity of the MYC response and rLon protease treatment

To better understand the consequences of MYC inhibition by Lon, the entire data set from infected mice (7 days) was compared between the placebo group and the Lon treated mice. Healthy uninfected mice were used as controls. By PCA (principal component analysis), the Lon treated mice formed a cluster near the healthy mice. The placebo group, in contrast, formed a cluster distant from the Lon treated group and the healthy mice (90% the variance, 1st principal component), suggesting a similarity in gene expression profile between healthy mice and Lon treated mice and inhibition of disease related gene expression.

To further investigate this hypothesis, a heat map was constructed based on the total number of detected genes (n=19863). A distinct set of up- or down-regulated genes was detected exclusively in the placebo group, suggesting a disease association. The LON treated mice mirrored the pattern of gene expression seen in healthy mice strongly supporting the conclusion that Lon treatment mainly inhibits the disease associated MYC response, comprising C-MYC, N-MYC and L-MYC.

Mechanisms of MYC inhibition by Lon and Lon domains

To investigate the mechanism of MYC inhibition by Lon, the full-length Lon protein and each globular domain was purified utilizing a C-terminal Strep-tag (Fig. 5a, b). Direct MYC degradation by Lon was quantified by mixing of rLon with recombinant c-MYC in the presence of ATP and Mg²⁺. A significant reduction in c-MYC protein levels was detected by Western blot analysis, after 120 min of incubation with rLon at 37 °C, compared to PBS (Fig. 5c). The Lon peptidase domain significantly reduced c-MYC protein levels, using the same protocol, and this effect was enhanced in the presence of the ATP binding AAA+ domain, which is required for full peptidase activity in bacteria²⁶ (Fig. 5d). Degradation was inactivated in the presence of protease inhibitors (Fig. 5e).

To investigate if the Lon protease can target intracellular MYC, human kidney carcinoma cells were treated with full-length rLon or each of the Lon domain peptides, using AZ647-fluorophore-labelling to enable visualization and quantitation. Cellular uptake of rLon was detected by live cell imaging (Fig. 5f) as significant rLon staining in the cytoplasm and nuclei. Each of the AZ647-labeled peptides showed a similar uptake pattern (Fig. 5f) also confirmed by Western blot analysis of whole-cell extracts as well as nuclear extracts of cells exposed to rLon or each of the peptides (Fig. 5g).

To quantify the effect of rLon treatment on cellular MYC protein levels, whole-cell extracts were subjected by Western blot analysis (Fig. 5h, i). c-MYC protein levels were significantly reduced in cells exposed to full-length rLon and this effect was especially pronounced in the nuclear extracts compared to cells treated with PBS. c-MYC levels were not affected by the individual peptide domains, however (Fig. 5h). L-MYC levels were reduced by rLon treatment, including isoform II (20 kDa) and isoform I (40 kDa) and strongly affected by the NTD, AAA+ and peptidase domains, although these isolated domains mainly affected the shorter isoform (Fig. 5h, i). N-MYC protein levels were significantly reduced by rLon treatment of cells treated with full-length rLon or the peptidase domain in whole-cell lysates and nuclear extracts (Fig. 5h).

Significant inhibition of MYC expression in rLon-treated cells was also investigated by qRT-PCR, using primers specific for MYC, MYCL (isoforms I and II) or MYCN (Fig. 5i). MYCL expression was strongly inhibited in human kidney cells exposed to the full-length protein (FC -9.7), compared to control cells and a further inhibition was detected in cells exposed to the NTD peptide (FC -38.4) and the peptidase domain peptide (FC -21), compared to PBS-treated control cells. In contrast, MYC expression was not significantly regulated by rLon or the Lon domain peptides (FC < 2) and the expression of MYCN was weakly regulated (FC -2.1 and -2.7), compared to PBS-treated control cells.

Targeting of MYC-MAX dimerization by Lon and the NTD domain peptide affecting MYC-dependent gene expression

The potent inhibition by of MYC-dependent gene expression in the disease models suggested that additional mechanisms might be involved. MYC protein domains that are targeted by rLon were therefore identified by in silico prediction technology (Fig. 6a-d). Using the AlphaFold multimer protocol, it was predicted that the Lon NTD peptide interacts with the C-terminal part of c-MYC (aa 356-439) and N-MYC (aa 354-464), with an iPTM score of 0.697 (Fig. 6a, b). The c-MYC and N-MYC sequences predicted to be targeted by the NTD of Lon correspond to the MAX binding site in the leucine zipper domain. Biacore analysis showed a dose-dependent binding of both full-length Lon protein and NTD peptide to His-tagged c-MYC (full length), c-MYC (C-terminal domain) and N-MYC (full length) with $K_D \sim \mu M$ range, showing affinity of full-length Lon protein and NTD peptide for both c-MYC and N-MYC (Fig. 6c). A structural overlay with the MYC-MAX complex crystal structure shows an almost perfect overlap of the NTD peptide with MAX (Fig. 6d), suggesting that NTD targets the MAX-binding site (Fig. 6d), a mechanism by which Lon displaces MAX and reduces MYC-MAX complex formation, which possibly explains the marked reduction in MYC-dependent gene expression by Lon treatment, observed in the in vivo models.

This hypothesis was addressed, using an ELISA-based TransAM c-MYC assay that quantifies c-MYC binding to its consensus E-box DNA sequence. The rLon protein and the NTD and

AAA+ domains reduced the binding of c-MYC to DNA (Fig. 6e). The effect of the NTD domain is consistent with the in silico prediction of NTD binding to the MAX binding site and the effect of the effect of the AAA peptide is consistent with the known affinity for DNA of this domain. Effects of Lon on MYC-dependent gene expression were further investigated, in
5 cells exposed to rLon or the NTD or AAA+ peptides. Genome-wide transcriptome analysis identified MYC targets, which were significantly inhibited in human kidney carcinoma cells treated with the NTD and AAA+ peptides, consistent with the in vitro data on MYC-MAX inhibition, with a weaker effect of full-length Lon (Fig. 6g, h).

The results identify the MYC family of transcription factors as targets for the bacterial Lon
10 protease, which is taken up by human cells and affects cellular MYC levels in several ways. First, degradation of MYC, executed by the peptidase domain together with the ATP binding AAA+ domain, was detected in vitro. Second, a reduction in cellular c-MYC, N-MYC and L-MYC levels following rLon treatment, was attributed to MYC degradation by the peptidase and AAA+ domains and the resulting inhibition of MYC expression. Finally, by targeting the
15 MAX-binding site, the Lon NTD domain was shown to inhibit MYC-MAX hetero-dimerization, which is essential for subsequent DNA binding of the helix-loop-helix domain that initiates MYC-dependent transcription²⁷.

Discussion

MYC deregulation is associated with severe disease states, most extensively defined for the
20 C-MYC proto-oncogene and its role in cancer. Selective inhibitors targeting deregulated MYC would ideally affect a majority of human cancers, but such inhibitors have been difficult to find. This study identifies Lon, a protein of bacterial origin, as a pan-MYC inhibitor, with potent effects that reverse disease-associated MYC deregulation in models of infection and cancer. Therapeutic effects were associated with broad inhibition of MYC dependent gene
25 networks, a loss of C-MYC, N-MYC and L-MYC overexpression. These effects are proposed here to reflect different rLon protein domains that either increase MYC degradation or inhibit the expression of MYC and MYC dependent genes, resulting in a reduction of pan-MYC levels. The results identify Lon as a new, powerful tool to neutralize the destructive effects of MYC deregulation in different disease states.

The rLon protease was previously shown to degrade C-MYC in vitro, in a time-dependent manner and to reduce c-MYC levels in human cells and cancer models. In this study, these effects are extended to include N-MYC and L-MYC, which showed a similar pattern of disease associated hyper-activation and sensitivity to rLon as C-MYC. Structure-function analysis suggested two mechanisms of inhibition. A synthetic peptide comprising the Lon
35 protease domain was shown to enter human carcinoma cells and reduce C-MYC, N-MYC and L-MYC levels in the cytoplasm and the nuclei. In addition, a synthetic peptide comprising the N-terminal domain was shown to translocate to the nuclei of human cells and act as a

potent transcriptional inhibitor, primarily of L-MYC, N-MYC and C-MYC expression. Thus, in addition to degrading C-MYC, N-MYC and L-MYC levels were reduced by the Lon peptidase domain. The effect of rLon was further extended to include transcriptional inhibition, executed by the N-terminal domain and mainly affecting L-MYC expression.

- 5 The MYC family of proteins share a basic helix-loop-helix leucine zipper structure with significant structural homology between conserved C terminal regions. The helix-loop-helix leucine zipper motif is essential for the DNA binding of the MYC transcription factors at sites of transcriptional initiation in many MYC-dependent promoters. In this study, the helix-loop-helix leucine zipper domain was identified as a target for the N-terminal domain of Lon.
- 10 Direct effects on MYC gene expression were detected but in addition, MYC-dependent gene networks were strongly inhibited, resulting in broad transcriptional inhibition. The efficient alignment of the Lon NTD with the helix-loop-helix leucine zipper motif, may contribute to this broad effect, by inhibiting the dimerization of MYC, which is required for DNA binding.

- MYC is ubiquitously expressed during tissue development and overexpressed in a broad
- 15 variety of tumors. MYCN is similarly regulated and deregulated in different cancer types and N-MYC can functionally replace C-MYC in the context of cellular growth and differentiation. Despite the structural similarity, they show distinct expression timings and tissue specificities. Although MYCL also plays roles in tumor and tissue development, it is reported that MYCL is essential for embryonic development and MYCL expression prime T cells in
 - 20 response to viral and bacterial infection. MYCL has more restricted pattern of tissue expression, which includes the kidneys. The findings identify these areas as targets of bacterial kidney infection, where high-resolution imaging identified abscesses containing bacteria and neutrophils in sham treated mice, but not in the rLon treated mice, consistent with a role of L-MYC hyper-activation for renal pathology and protection by rLON against
 - 25 renal toxicity and tissue damage at these sites.

- Many approaches to MYC inhibition have proven unsuccessful, due to the absence of a suitable pocket for binding of low molecular weight inhibitors and the lack of catalytic activity. Furthermore, its nuclear localization has made MYC more difficult to target with therapies such as monoclonal antibodies. The full-length Lon protease and the peptidase
- 30 and NTD domains were shown to reach tumor cell nuclei and to efficiently reduce nuclear C-MYC, N-MYC and L-MYC protein levels. Approaches to MYC inhibition further include repressing transcription and translation by targeting the obligate MYC interaction partner MAX, modulating MYC stability or targeting pathways upstream of MYC or blocking downstream genes exhibiting synthetic lethality with overexpression of MYC. ref This
 - 35 mechanism may not be relevant for rLon, as the rLon protease has shown selectivity for c-MYC, and a lack of degradation of the MYC partner MAX, CEBP- δ , AKT, CDK9, CPEB1, GRP75, Actin or histone H3⁸. The OmoMYC mini-protein, corresponding to the HBLH domain

of MYC, acts by blocking the binding of all 3 forms of MYC to their target promoters and has been shown to exhibit anti-cancer activity in a range of preclinical models, with minimal side effects. OmoMYC is thought to direct MYC to a particular subset of target genes whose repression promotes apoptosis by interfering with MYC/MAX binding to the E-box. The Lon protease has a more general effect on gene expression, preferentially inhibiting the disease related deregulation of MYC. Furthermore, the Lon protease was shown to strongly inhibit L-MYC expression, potentially as a result of the degradation of C-MYC and N-MYC and the resulting inhibition of L-MYC expression.

Disease specific therapeutic interventions should ideally maintain health, without adversely affecting tissue homeostasis. In this study, rLon treated healthy mice did not show evidence of disease or major changes in MYC expression, suggesting that Lon may not target normal MYC expression to the same extent as the disease-associated MYC response. The molecular basis of this apparent specificity remains unclear, but several scenarios may be discussed. First, the rLon protease may target disease specific modifications of C-MYC, that optimize Lon binding and the degradation of its target. MYC dysregulation during cancer and infection may create structural changes such as specific phosphorylation patterns that enhance the binding of the peptide domain and facilitate MYC degradation. Second, MYC overexpression may potentially enhance the presentation of the target or its cellular localization, making it more available for the protease to execute its proteolytic effect. Third, the access of rLon or the N-terminal domain to specific MYC promoters, may be enhanced in the open chromatin of rapidly dividing cancer cells, possibly giving rise to the pronounced transcriptional inhibition, which was detected in disease but not present in healthy tissues.

Example 2

Methods

25 Expression and purification of recombinant Lon protease

The nucleotide sequences of different domains of Lon protease, including the Nterminal domain (NTD), the ATPase domain (AAA+), the peptidase domain (Pep.), as well as the Lon full-length protein, were chemically synthesized (Genscript). These genes were cloned and tagged with Strep-tag II using pASK-63A+ vectors (IBA), carrying the β -lactamase gene resulting in ampicillin resistance as a selectable marker, along with a tetracycline promoter/operator system that tightly controls their expression. Then, four different pASK-63A+ expression constructs containing different Lon protease domains were individually transformed into competent *E. coli* BL21 (DE3) competent cells using the heat-shock method. Selected single colonies were then inoculated in Luria-Bertani (LB) media containing 100 mg/ml ampicillin and incubated at 37 °C with shaking at 200 rpm until the optical density at 600 nm (OD600) reached about 0.5. The cell pellet was then collected by

centrifugation at 14,000 x g for 10 minutes, and the resulting pellets were disrupted by lysis buffer (50 mM Tris pH 7.4, 100 mM NaCl) by sonication on ice (20 second bursts with a 40-second pause between each burst, approximately 10 minutes), followed by centrifugation (46,000 x g, 30 minutes, 4 °C). Purification was performed using the poly-prep chromatography column (Bio-Rad) with a diameter of 0.8 cm packed with Strep-Tactin Sepharose (IBA

Lifesciences), according to manufacturer's instructions. The Lon protease was then collected from the column by using elution buffer (50 mM Tris pH 7.4, 100 mM NaCl, 5 mM desthiobiotin, pH 7.4) and dialyzed with the Snakeskin Pleated Dialysis Tubing (10,000 MWCO) against the buffer containing 50 mM Tris pH 7.4, 100 mM NaCl, 10% glycerol by slow shaking at room temperature for at least one hour followed by overnight at 4 °C. The eluted fractions were used either immediately or stored at -80 °C.

***In silico* prediction of MYC-Lon interactions**

The Lon sequence used in the AlphaFold multimer protocol⁴³ was identical to that used for recombinantly produced globular domains. The MYC sequences used were as follows: UniProt ID: P01106-1 (c-MYC), P04198 (N-MYC), and P12524-1 (L-MYC). These sequences used were either full-length or divided into overlapping fragments to ensure that the flexible nature of MYC proteins would not shield potential binding sites of global Lon domains. Sequences were then submitted to CoLabFold using the AlphaFold2 multimer protocol. For searches using sequence fragments, all predictions were subjected to an initial visual sanity check (Pymol open source), and complexes were filtered out by either the Lon domain purification Strep-tag or those MYC termini arising from dividing the sequence. Hits were then sorted according to their respective iPTM scores.

Surface plasmon resonance analysis

Real time biomolecular interaction analysis was performed by surface plasmon resonance (SPR) technology using a Biacore X100 instrument (Cytiva Life Sciences, Uppsala, Sweden) loaded with a Sensor Chip NTA (BR100034, Cytiva Life Sciences).

All experiments were carried out at 25 °C with a constant flow rate of 30 µl/ minute with 0.1 M HEPES, 1.5 M NaCl, 0.03 M EDTA and 0.5% v/v Surfactant P20 (BR100826, Cytiva Life Sciences). The cell without nickel loading served as a reference for nonspecific binding of proteins and chemicals. Biacore X100 control software was programmed to first prime the instrument with running buffer (0.1 M HEPES, 1.5 M NaCl, 0.03 M EDTA and 0.5% v/v Surfactant P20). NiCl₂ were flowed into saturate the NTA with nickel ions followed by a 3 mM EDTA wash. The Histidine-tagged ligand was prepared in running buffer at concentration ~0.2 µM and injected over the nickel activated sensor surface to capture the

His-tagged ligand, full-length c-MYC (CSBEP015270HU, Cusabio), C-terminal c-MYC (TP762404, Origene) or full-length N-MYC (AP72694, Signalway). Sample was injected over the captured His-tagged ligand with five different dilutions for monitoring the interaction of sample with His-tagged ligand. Finally, the chip was regenerated with injection of 350 mM EDTA. Binding affinities were calculated from rate constants obtained from fitting the data to a 1:1 Langmuir binding model.

DNA-binding ELISA

The TransAM c-MYC Kit (43396, Active Motif) was used to assess whether c-MYC activation was inhibited by Lon. Recombinant c-MYC/MAX complex (81087, Active Motif) was used as a standard. Recombinant Lon and individual domains were added to 5 or 2.5 ng c-MYC/MAX complex, and DNA-bound MYC was quantified according to the manufacturer's instructions.

Cell culture

A498 human kidney carcinoma cells (ATCC, HTB-44) were cultured in RPMI-1640 supplemented with 1 mM sodium pyruvate, 1 mM nonessential amino acids and 10% fetal bovine serum (FBS) overnight at 37 °C with 90% humidity and 5% CO₂ in 6-well plates (for western blots) or 8-well chamber slides (for confocal imaging). Cells were treated in fresh, serum-free supplemented RPMI-1640 medium.

Cell treatment with rLon and the rLon domain peptides

Human kidney carcinoma cells were grown overnight in either 8-well chamber slides (Lab-Tek II chambered no. 1.5 German cover-glass system; Thermo Fisher Scientific), six-well plates or 10-cm-diameter dishes in appropriate media supplemented with fetal calf serum (FCS). Cells were washed and treated with 902 nM recombinant protein in serum-free medium for 3.5 hours.

Treated cells were analyzed by quantitative real-time (qRT-PCR), whole-genome transcriptome analysis, confocal microscopy (magnification 63x), immunoblotting or cell toxicity assays.

Cellular uptake

The N-terminus of recombinant proteins were labeled with AZ647 using the NHS functional group according to the manufacturer's instructions (A10277, Thermo Scientific). Free dye was removed using a dialysis cassette (D-Tube™ Dialyzer Midi) followed by centrifugation (3,500 x g, 40 minutes, 4 °C) with an Amicon centrifugal filter unit (Amicon Ultra15). The concentration of the labeled complex was determined using a Cary UV-Visible spectrophotometer or NanoDrop 2000c (Thermo Scientific). For livecell uptake, cells were treated with 902 nM or 6.3 μM recombinant protein for three hours. Nuclei were

counterstained with Hoechst 33342 (1:2000, Molecular Probes, Cat# H1399), imaged by laser-scanning confocal microscopy (LSM900 confocal microscope, Carl Zeiss) and quantified by ImageJ software (NIH).

Real-time quantitative polymerase chain reaction

- 5 Total RNA was extracted from kidney epithelial cells with the RNeasy Mini Kit (Qiagen) after disruption in RLT buffer supplemented with β -mercaptoethanol (1%) using Qi shredder (Qiagen). Total RNA was reverse transcribed using the SuperScript III First-Strand Synthesis System for PCR (Thermo Fisher Scientific) and quantified in real time using QuantiTect SYBR Green-based PCR Kits for mRNA (Qiagen) on a
- 10 RotorGene Q cycler (Qiagen). Human-specific primers, including QuantiTect Primer Assay Hs_MYCN_1_SG (QT00201404), Hs_MYCL_1_SG (QT01004710), Hs_MYCL_va.1_SG (QT01004703), Hs_MYC_1_SG (QT00035406), Hs_IRF7_1_SG (QT00210595) and Hs_GAPDH_1_SG (QT00079247) were used; all primers were from ag. qRT-PCRs were run in duplicate, and the $2^{-\Delta\Delta C_t}$ method was used to calculate relative gene expression.

15 Whole-genome transcriptome analysis of kidney cells

- Total RNA was extracted from kidney epithelial cells with the RNeasy Mini Kit (Qiagen) after disruption in RLT buffer supplemented with β -mercaptoethanol (1%) using Qiashredder (Qiagen). Exactly 100 ng of total RNA was amplified and fragmented using the Affymetrix WT PLUS Reagent, followed by hybridization onto Clariom S Human arrays (Affymetrix) for
- 20 16 hours at 45 °C. Then, the sample was washed, stained, and scanned using the GeneChip System as per manufacturer's instruction (Affymetrix). Data were normalized using Robust Multi Average implemented in Transcriptome Analysis Console software (v.4.0.1.36, Applied Biosystems, Thermo Fisher Scientific). Relative expression was analyzed by ANOVA using the empirical Bayes method, and genes with an absolute fold change > 1.5 were considered
- 25 differentially expressed. Heatmap was constructed in GraphPad Prism 10.

Western blot analysis

- Cells were lysed with NP-40 lysis buffer supplemented with protease and phosphatase inhibitors (Roche Diagnostics). For nuclear extract preparation, cell compartments were fractionated using a Qproteome Cell Compartment kit (Qiagen). Proteins were separated by
- 30 SDS-polyacrylamide gel electrophoresis (SDS-PAGE; 4-12% Bis-Tris gels, Invitrogen) and blotted onto polyvinylidene difluoride membranes (Bio-Rad), which were then blocked and stained with primary rabbit anti-Lon (1:2,000 in 5% BSA/PBS; Biorbyt, orb231326), rabbit anti-c-MYC (1:1,000 in 5% BSA/PBS-T; Abcam, no. ab3968), rabbit anti-L-Myc (Aviva Systems Biology, ARP35770) or rabbit anti-NMyc (Cell Signaling Technology, 51705)
- 35 antibodies overnight at 4 °C. Membranes were then washed, stained with goat anti-rabbit

IgG horseradish peroxidase-conjugated secondary antibodies (1:4,000 in 5% BSA/PBS-T, 1 hour at room temperature), and washed again, imaged using enhanced chemiluminescence plus detection reagent (GE Healthcare), and protein signals were quantified by ImageJ. β -Actin (1:4,000 in PBS-T, 1 hour at room temperature; Sigma, A5316), GAPDH (1:4,000 in PBS-T, 1 hour at room temperature; Santa Cruz Biotechnology, sc-25778) or histone H3 (1:4,000 in PBS-T, 1 hour at room temperature; Abcam, ab1791) was used as a loading control.

Cell toxicity assays

Cells (4×10^4 cells per well) were seeded in serum-free RPMI-1640 on 96-well plates (Tecan Group). Cells were treated with rLon protease or different domain peptides and incubated for three hours. Cell viability or death was quantified by two different assays. The released ATP was determined using the luminescence-based ATP Lite kit (Perkin Elmer) and mitochondrial enzyme activity was determined by the Presto Blue assay (Invitrogen). Assays were done in triplicate, and mean values were calculated.

15 *In vitro* degradation assay

Recombinant c-MYC (CSB-EP015270HU, Cusabio) or N-MYC (AP72694, SAB Signalway Antibody) protein (0.2 μ g) was incubated with 0.2-1 μ g of rLon protease, 0.5 μ g of AAA+ or peptidase domain peptides (produced in-house, either with or without protease inhibitors) or PBS in a reaction buffer consisting of 20 mM MgCl_2 in the presence of ATP (2 mM) in 50 mM Tris pH 8.0. Proteins were incubated at 37 °C for

120 minutes and denatured using 10 mM dithiothreitol (DTT) in NuPAGE LDS Sample Buffer (Thermo Fisher Scientific) at 100 °C for five minutes before immunoblot analysis using a c-MYC or N-MYC antibody (Abcam, ab39688).

Bacterial strains

25 The prototype APN strain *E. coli* CFT073 (O6:K2:H1)⁴⁴ was cultured on tryptic soy agar (TSA) plates at 37 °C for 16 hours, harvested in phosphate-buffered saline (PBS, pH 7.2) and diluted as appropriate¹⁸.

To quantify effects of Lon on bacterial viability, *E. coli* CFT073 bacteria (CFU = 10^4) were first cultured in 1000 μ l LB media for 3.5 hours at 37 °C after which Lon (250 or 500 μ g/ml) or cefotaxime (100 mg/ml) was added to the bacterial culture. The OD600 (optical density at 600 nm) was measured at different timepoints (0, 1, 2, 3, 4 and 5 hours).

Experimental kidney infection model

Irf3^{-/-} mice on a C57BL/6 background were kindly provided by the Riken Bioresource Center, Japan, with permission from T. Taniguchi⁴⁵. Mice were bred at Lund University,

BMC animal facility, Lund, Sweden, and housed in specific pathogen-free individually ventilated cages (IVCs) at a constant temperature of 23 °C on a 12-hour light-dark cycle with lights on at 7:00 a.m. and *ad libitum* access to food and water. Female *Irf3*^{-/-} mice were used for experiments at 9-15 weeks of age.

Mice were anesthetized by intraperitoneal (*i.p.*) injection of a cocktail of ketamine (1.48 mg in 100 µl of 0.9% NaCl solution, Intervet) and xylazine (0.22 mg in 100 µl of 0.9% NaCl solution, Vetmedic) and infected by intravesical inoculation with *E. coli* CFT073

(10⁸ CFU in 0.05 mL)⁴⁶ through a soft polyethylene catheter (2x10⁹ CFU/ml, 50 µl). Urine samples were obtained from each mouse by holding the animal over a sterile 1.5 ml Eppendorf tube before infection and at regular times after infection (24 hours, 3 days, 5 days and 7 days). The urine drops were immediately collected, viable bacteria counts were determined by growth on TSA plates (37°C, overnight) after serial dilutions as appropriate. Neutrophils in uncentrifuged urine were counted using a hemocytometer chamber. The mice were sacrificed by using an overdose of isoflurane anesthesia (Dechra, Cat# 200-129) after 24 hours and seven days. Organs were aseptically removed, and macroscopic pathology was documented by photography. Tissue samples were fixed for 24 hours in 4% paraformaldehyde (PFA) at 4 °C and embedded in paraffin for H&E and immunohistochemistry analysis. The degrees of hyperemia, edema and abscesses were scored on a scale from 0-10, where 0 was similar to that of uninfected controls, and 10 indicated the most severe hyperemia, edema and abscesses. One kidney per mouse was divided for RNA extraction (on dry ice) and immunostaining (in O.C.T. compound or 10% (v/v) formalin).

Lon protease treatment in mice

Recombinant bacterial Lon protease was administered intraperitoneally (rLon, 100 µl, 250 µg/ml) to *Irf3*^{-/-} mice once a day for eight days starting one day before intravesical infection with *E. coli* CFT073 (day 0). Placebo-treated mice received intraperitoneally 100 µl of PBS with the same schedule as Lon-treated mice. The mice were sacrificed by using an overdose of isoflurane anesthesia (Dechra, Cat# 200-129) after 24 hours and seven days.

Apc^{Min/+} (C57BL/6J-*Apc*^{Min}/J, RRID:IMSR_JAX002020)^{47,48} mice and C57BL/6 mice were obtained from Jackson Laboratories at approximately eight weeks of age. *Apc*^{Min/+} mice carry a heterozygous mutation in the murine *Apc* (*adenomatous polyposis coli*) locus, which consequently encodes a nonsense mutation at codon 850 resulting in multiple intestinal

neoplasia (*Min*). C57BL/6J-*Apc^{Min}*/J mice, which develop spontaneous intestinal adenomas, were bred at Lund University and used at 11 weeks old. Mice were treated twice daily for 14 days by oral gavage of rLon protease (200 µl, 250 µg/ml). At sacrifice, intestines were aseptically harvested, and segments of the intestines were prepared for mRNA extraction.

5 Whole-genome transcriptome analysis of mouse tissues

- Total RNA was extracted from murine kidneys or intestines with the RNeasy Mini Kit (Qiagen) after disruption in RLT buffer supplemented with β-mercaptoethanol (1%) using Precellys Lysing kits (Bertin Technologies) and TissueLyser (Qiagen). For treated cells, RNA was extracted with the RNeasy Mini Kit and Qias shredder (Qiagen). Exactly 100 ng of total
- 10 RNA was amplified and fragmented using the Affymetrix WT PLUS Reagent or GeneChip 3' IVT Express kits, followed by hybridization onto Clariom S Mouse HT or Mouse Genome 430 PM (Affymetrix) for 16 hours at 45 °C. Then, the sample was washed, stained using the Affymetrix fluidic station 450 as per manufacturer's instruction (Hybridization, Washing and Staining kit, Affymetrix).
- 15 Microarrays were immediately scanned using the GeneTitan or GeneAtlas system (Affymetrix). Data were normalized using Robust Multi Average implemented in Transcriptome Analysis Console software (v.4.0.1.36, Applied Biosystems, Thermo Fisher Scientific). Relative expression was analyzed by ANOVA using the empirical Bayes method, and genes with an absolute fold change ≥ 2.0 were considered differentially expressed.
- 20 Differentially expressed genes were functionally analyzed using Ingenuity Pathway Analysis software (IPA, Qiagen), and data are shown in heatmaps constructed in GraphPad Prism 9.

Immunohistochemistry

- Tissues embedded and frozen in OCT were cryosectioned (8 µm thick sections, Leica microtome), collected on positively charged microscope slides (Superfrost/Plus; Thermo
- 25 Fisher Scientific), fixed in acetone-methanol (1:1, 10 minutes), permeabilized (0.2% Triton X-100, 5% normal goat serum/PBS) and incubated overnight at 4 °C with primary rabbit monoclonal anti-c-Myc (1:100; Abcam, ab32072), rabbit anti-L-Myc (1:100; Aviva Systems Biology, ARP35770) or rabbit anti-N-Myc (1:100; Cell Signaling Technology, 51705) antibodies, followed by incubation for one hour at room temperature with Alexa 488- or
- 30 Alexa 568-labeled rabbit anti-rat and goat anti-rabbit IgG secondary antibodies (1:100; Molecular Probes, A-21210, A-11001 and A-11011). Nuclei were counterstained with DAPI (4',6-diamidino-2-phenylindole; 0.05 mM; Sigma-Aldrich) for 15 minutes at room temperature. Slides were examined by fluorescence microscopy (AX60, Olympus Optical). Richard-Allan Scientific Signature Series Hematoxylin 7211 and Eosin-Y 7111 (Thermo
- 35 Scientific) were used to counterstain the tissue sections.

Quantification and statistical analysis

For data that had a Gaussian distribution (D'Agostino & Pearson normality test), significant differences were determined by two-tailed *t* test, one-way ANOVA with by Dunn's correction or two-way ANOVA with Bonferroni's multiple comparison. For nonparametric data, either
5 the two-tailed Mann-Whitney U test or Kruskal-Wallis test with Dunn's correction was used, and paired data were analyzed by the Wilcoxon matched-pairs signed-rank test. Statistical significance was determined by GraphPad (Prism v. 9), and significance was assigned at * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$. All data are presented as the median or mean, and all error bars represent the standard deviation (s.d.) or standard error mean (s.e.m.).
10 Sample size (*n*) indicates either the number of independent experiments or the number of mice used in the study.

Results

Identifying MYC targets of by Lon and the Lon peptides

To investigate the mechanism of MYC inhibition by Lon, the full-length Lon protein and the
15 N-terminal domain (NTD), the hexameric AAA+ domain and the C-terminal peptidase domain peptides were recombinantly expressed and purified utilizing a Cterminal Strep-tag, as shown by Western blot analysis (Fig. 7a, b).

In silico prediction technology was used to identify MYC protein domains targeted by rLon and the Lon peptide domains (Fig. 7c-e). The AlphaFold multimer protocol prediction
20 indicated a low score for the interaction between full-length c-MYC and fulllength Lon (Fig. 13). However, the resulting model revealed the N terminal domain of Lon positioned near the C-terminal domain of c-MYC, prompting further analysis into the interactions between these two domains. It was predicted that the Lon NTD peptide interacts with the C-terminal part of c-MYC (aa 356-439) and NMYC (aa 354-464), with an iPTM score of 0.697 (Fig. 7c,
25 d) but not with L-MYC (iPTM score of 0.246 Fig. 7e). Weak interactions were detected between the full-length MYC proteins and the Lon peptidase domain.

Surface plasmon resonance (SPR) analysis, through Biacore technology, was used to quantify the binding of the rLon protein and the Lon peptides to full-length c-MYC, NMYC and the C-terminal domain, which is highly conserved. The His-tagged full-length c-MYC or
30 N-MYC were captured on NTA Sensor chips followed by a dose-dependent binding analysis with rLon protein or the Lon peptides. rLon and the NTD peptide were shown to bind to His-tagged c-MYC and N-MYC, in a dose-dependent manner, with a K_D in the ~30 to 40 μ M range. Furthermore, the NTD peptide bound to the His-tagged C-terminal domain of c-MYC (Fig. 7f). The SPR analysis further detected the binding of the peptidase- and AAA+
35 domains to full-length c-MYC. The binding affinity of rLon and its domains is similar to that reported for MYCMI-6, a small molecule inhibitor of Myc-driven transcription, which binds to

the MYC bHLHZip domain²³. The apparent discrepancy between the lack of *in silico* prediction for the peptidase domain and the positive SPR analysis, confirmed that AlphaFold cannot clearly provide folding patterns within intrinsically disordered regions, which are usually the targets for degradation²⁴. Due to the unavailability of full-length L-MYC protein, the binding interactions of full-length Lon and Lon peptides with L-MYC could not be determined through the SPR analysis.

A structural overlay with the MYC-MAX complex crystal structure showed an almost perfect overlap of the NTD peptide with MAX bound to MYC (Fig. 7g), suggesting that NTD targets the MAX-binding site, a mechanism by which Lon would displace MAX and reduce MYC-MAX complex formation. This hypothesis was addressed, using an ELISA-based TransAM c-MYC assay that quantifies c-MYC binding to its consensus E-box DNA sequence. The rLon protein and the NTD and AAA+ domains reduced the binding of c-MYC to DNA (Fig. 7h). The effect of the NTD domain is consistent with the *in silico* prediction of NTD binding to the MAX binding site, and the effect of the AAA peptide is consistent with the known affinity for DNA of this domain²⁵.

Cellular uptake of the Lon and Lon peptide domains and effects on cellular MYC levels

The Lon protease was shown to enter human cells and reduce intracellular MYC levels, but the cellular interactions of the Lon peptides has not previously been investigated. To address this question, human kidney carcinoma cells were treated with each of the Lon peptides, using AZ647-fluorophore-labelling to enable visualization and quantification. Each of the AZ647-labeled peptides showed a similar pattern of cellular uptake (Fig. 8a) with significant staining in the cytoplasm and nuclei, also confirmed by Western blot analysis of whole-cell extracts of cells exposed to rLon or each of the peptides (Fig. 8b; Fig. 9) with a complex pattern suggesting fragmentation of rLon creating peptide domains.

To quantify the overall effects on cellular c-MYC, N-MYC and L-MYC protein levels, human kidney cells treated with rLon or each of the Lon peptides were fixed and imaged by confocal microscopy, after staining with anti-c-MYC-, N-MYC- or L-MYC- specific antibodies. c-MYC, N-MYC and L-MYC showed a similar pattern of reduction in staining after rLon and NTD treatment compared to control cells (Fig. 14c-e). In contrast, cells treated with the peptidase domain did not show a significant reduction in c-MYC or N-MYC signal but redistribution was detected within the cells, from the cytoplasm to the nuclei. The cytoplasmic c-MYC signal was markedly attenuated, whereas the nuclear and perinuclear c-MYC signal was enhanced, suggesting c-MYC translocation and/or increased activation (Fig. 8c, d). Furthermore, rLon and the NTD peptide significantly reduced L-MYC levels. However, unlike the other MYC members, the L-MYC signal was also reduced in peptidase-treated

cells compared to control (Fig. 8e). The AAA+ domain peptide decreased the N-MYC level but increased c-MYC and L-MYC levels.

The results confirm that rLon treatment reduces cellular MYC levels and identifies NTD as a potent C-, N- and L-MYC inhibitor in treated cells.

5 Mechanism of inhibition of MYC gene expression by Lon and Lon domain peptides

The inhibition of MYC/MAX complex formation by Lon and the NTD peptide, predicted that gene expression would be affected in rLon- or NTD-treated cells. *MYC* expression was therefore quantified by qRT-PCR, using primers specific for *MYC*, *MYCN* or *MYCL* (isoforms I and II) (Fig. 8f). *MYC* expression was not significantly regulated by rLon or the Lon domain peptides (FC < 2) and the expression of *MYCN* was weakly regulated (FC -2.1 and -2.7), compared to PBS-treated control cells. In contrast, *MYCL* expression was strongly inhibited in human kidney cells exposed to rLon (FC -9.7), compared to control cells and a further increase in inhibition was detected in cells exposed to the NTD peptide (FC -38.4) or the peptidase domain peptide (FC -21), compared to control.

Effects of Lon on MYC-dependent gene expression were further investigated in cells exposed to rLon, NTD, peptidase or AAA+ peptides. Genome-wide transcriptome analysis identified MYC targets, which were significantly inhibited in human kidney carcinoma cells treated with the NTD and AAA+ peptides, with a weaker effect of rLon and the peptidase domain, consistent with the *in vitro* data on MYC-MAX inhibition (Fig. 8g, h). Inhibited genes included zinc finger proteins (ZFPs), indicating a mechanism contributing to the effect on the transcriptional machinery.

Mechanisms of MYC degradation by Lon and Lon peptide domains

Direct MYC degradation by Lon was quantified *in vitro*, by mixing of rLon with recombinant c-MYC or N-MYC in the presence of ATP and Mg²⁺. A significant reduction in c-MYC protein levels was detected by Coomassie staining and Western blot analysis, after 120 min of incubation with rLon at 37 °C, compared to PBS (Fig. 9a). The Lon peptidase domain significantly reduced c-MYC and N-MYC protein levels, using the same protocol, and this effect was enhanced in the presence of the ATP binding AAA+ domain, which is required for full peptidase activity in bacteria²⁶ (Fig. 9b). Degradation was inactivated in the presence of protease inhibitors (Fig. 9c). Full-length L-MYC was not available for *in vitro* analysis.

To quantify the effect of rLon treatment on cellular MYC protein levels, whole-cell extracts were subjected to Western blot analysis (Fig. 9d-f). c-MYC protein levels were significantly reduced in cells exposed to full-length rLon and this effect was especially pronounced in the nuclear extracts compared to cells treated with PBS. c-MYC levels were not affected by the individual peptide domains, however (Fig. 9d-f). N-MYC protein levels were significantly

reduced by rLon treatment of cells treated with full-length rLon or the peptidase domain in whole-cell lysates and nuclear extracts (Fig. 9d-f). L-MYC levels were more strongly reduced by rLon treatment than c-MYC and N-MYC, including isoform II (20 kDa) and isoform I (40 kDa) and strongly affected by the NTD, AAA+ and peptidase domains, although these
 5 isolated domains mainly affected the shorter isoform (Fig. 9d-f).

The results identify the MYC family of transcription factors as targets for the bacterial Lon protease, which is taken up by human cells and affects cellular MYC levels in several ways. First, *in silico* analysis predicted that the MAX-binding site is targeted by the Lon NTD domain, which was shown to inhibit MYC-MAX hetero-dimerization, which is essential for
 10 subsequent DNA binding of the helix-loop-helix domain that initiates MYC-dependent transcription²⁷. Second, degradation of MYC, executed by the peptidase domain together with the ATP binding AAA+ domain, was demonstrated *in vitro* suggesting that the reduction in cellular c-MYC, N-MYC and L-MYC levels following rLon treatment, can be attributed to both mechanisms. Interestingly, the regulation of L-Myc was attributed to the
 15 transcriptional effect and the resulting loss of c-MYC and N-MYC activation of L-MYC expression, rather than to degradation by the peptidase and AAA+ domains.

The MYC response to infection is inhibited by Lon treatment in disease-prone mice

Mice carrying *Irf3* deletions are susceptible to renal infections and *Irf3*^{-/-} mice are a relevant model of human disease^{28,29}. The MYC response to renal infection was therefore investigated
 20 in *Irf3*^{-/-} mice, infected by intravesical injection of the uropathogenic *E. coli* strain CFT073 (Fig. 10a). Total renal RNA harvested from infected mice or uninfected mice was subjected to genome-wide transcriptomic analysis and disease-associated genes were identified after 24 hours and seven days, by comparing the infected mice to uninfected controls.

The expression of *Myc* and that of MYC-dependent genes increased in response to infection
 25 (486 significantly regulated MYC-related genes at 24 hours and 786 after seven days compared to uninfected controls) and the majority of those genes were activated (Fig. 4b, c). The increase in total gene expression showed similar kinetics as the MYC-dependent genes, compared to uninfected controls. Analysis of renal tissues further showed an increase in renal c-Myc protein levels in infected *Irf3*^{-/-} mice, compared to uninfected
 30 controls, specifically along the mucosal lining of the renal pelvis and in the renal papillae, which are primary infection sites (Fig. 10e, f).

We next investigated whether Lon acts as a MYC inhibitor in this model. *Irf3*^{-/-} mice were intraperitoneally injected with recombinant Lon protease (rLon, 250 µg ml⁻¹) once a day for eight days starting one day before intravesical infection with *E. coli* CFT073 (Fig. 10a). A
 35 significant decrease in renal *Myc* gene expression was detected in the rLon-treated group (2 significantly regulated genes compared to 786 in the placebo group, on day 7, Fig. 10c, d).

Analysis of renal tissues revealed a decrease in c-Myc staining in rLon-treated infected *Irf3*^{-/-} mice compared to the placebo group, confirming the treatment effect at the protein level (Fig. 10e, f).

Functional analysis of the MYC-dependent genes identified innate immunity as the most strongly regulated gene category in the placebo group (Fig. 4g), with activation of c-MYC-dependent regulators of innate immunity; toll-like receptors, IL-1 β , tumor necrosis factor (TNF), type I interferons, as well as calcium binding proteins (S100A8, S100A9), which activate neutrophil recruitment (Fig. 4g). The MYC-dependent innate immune response was virtually abolished by rLon treatment (Fig. 4h). Major innate immune signaling pathways, were no longer regulated in rLon-treated mice, including the interferon, toll-like receptor, NF- κ B and inflammasome signaling pathways (Fig. 10i). The results suggest that infection triggers c-MYC deregulation and an exaggerated innate immune response, which is inhibited or virtually abolished, in rLon-treated *Irf3*^{-/-} mice, compared to the placebo group.

rLon treatment inhibits the activation of L-MYC and N-MYC in models of infection and cancer

To investigate if the effect of rLon extends to additional MYC family members, the RNA data set was analyzed specifically for L-MYC and N-MYC dependent gene expression (Fig. 11a-d). The expression of L-Myc and N-Myc was strongly increased at 24 hours and seven days after infection compared to that in uninfected control mice. Approximately 25% of genes annotated as related to N-Myc and 33% of genes annotated as related to L-Myc showed increased expression. L-Myc and N-Myc staining was further increased along the renal pelvis and in the papillae, confirming the N-Myc and L-Myc response at the protein level in infected *Irf3*^{-/-} mice (7 days, Fig. 11e, f).

These L-Myc and N-Myc responses to infection were virtually abolished by rLon treatment in *Irf3*^{-/-} mice at both the 24-hour and seven-day time points (Fig. 11a-f). No L-MYC-dependent gene and only two N-MYC-related genes were differentially expressed in rLon-treated mice, seven days after infection (Fig. 11b, d) and L-Myc and N-Myc staining was inhibited in tissues from infected, rLon treated *Irf3*^{-/-} mice (Fig. 5e, f). These results suggest that N-Myc and L-Myc are activated by infection in *Irf3*^{-/-} mice and that rLon inhibits N-MYC and L-MYC deregulation in infected kidneys.

To address if the effects of rLon on L-MYC and N-MYC expression translates to other disease states, RNA data sets from rLon-treated *Apc*^{Min/+} mice were interrogated for changes in N-MYC- and L-MYC-dependent gene expression, compared to the placebo group (Fig. 11g-m). N-Myc and L-Myc expression was significantly activated in tumorbearing *Apc*^{Min/+} mice in the placebo group (Fig. 11h-j) but not in the rLon treatment group (Fig. 11k-m). These results

suggest that N-Myc and L-Myc are deregulated in the intestines of *Apc^{Min/+}* mice and that rLon acts as a pan-MYC inhibitor in these models of bacterial infection and cancer.

rLon-treated mice are protected from renal toxicity

Acute pyelonephritis is an important cause of acute destruction of renal tissues and lasting renal damage³⁰. We therefore investigated if genes that regulate renal toxicity are affected by rLon treatment. A significant number of renal damage-associated genes were regulated in the placebo group compared to the uninfected control group and 93/271 of those genes were defined as MYC-related (34%, Fig. 12a, b). In contrast, only one gene associated with renal toxicity was significantly regulated in the rLon-treated group, and there was no evidence of a broader MYC-dependent renal damage response (Fig. 12a-d). The results suggest that rLon targets and inhibits genes involved in renal toxicity, renal injury or kidney-specific cell death, which drive renal damage.

The results described above, indicated that Lon might target disease-associated MYC deregulation more strongly than physiological MYC expression and tissue homeostasis. This question was addressed by quantifying homeostatic MYC dependent gene expression in healthy, uninfected *Irf3^{-/-}* mice, compared to rLon- or placebo-treated infected *Irf3^{-/-}* mice. Gene expression unrelated to disease was not significantly affected by rLon treatment and treated mice remained similar to the uninfected mice. Principal component analysis of the RNA data set showed that the rLon-treated group clustered adjacent to that of uninfected mice, while data for the placebo group clustered at a significant distance from the healthy mice (90% variance, 1st principal component, Fig. 12e).

Volcano plots were further constructed to visualize the overall difference in gene expression between uninfected *Irf3^{-/-}* mice, rLon-treated mice and placebo-treated mice, including genes that were expressed but not differentially regulated by infection and effect of Lon treatment on these gene categories (Fig. 12f-h). The disease response, comprised a total of 3,695 genes, including 786 MYC-related genes, which were differentially expressed in the placebo group compared to uninfected *Irf3^{-/-}* mice (Fig. 12f, g). In contrast, very few of these genes were differentially expressed in the rLon-treated mice ($n = 43$) compared to uninfected *Irf3^{-/-}* mice (Fig. 12f, g), illustrating the targeting by rLon of the disease-associated increase in gene expression that occurred in the placebo group.

These results indicate that Lon mainly inhibits MYC deregulation caused by infection, without causing major disruptions of healthy MYC gene expression machinery.

Discussion

This study investigated the mechanism of MYC inhibition by Lon and identified two parallel mechanisms of action; one in which NTD blocks the c-MYC and N-MYC-MAX complex

formation, thereby interfering with MYC-driven gene expression and another in which the peptidase domain, enhanced by the AAA+ domain degrades the MYC proteins. No off-target effects were detected suggesting that rLon treatment mainly affects the deregulation of MYC in these disease models. Together, these results introduce Lon protease as a new, versatile tool that could be used to neutralize the destructive effects of MYC deregulation in different disease states.

Of the MYC family, c-MYC is the most investigated. This short-lived protein is degraded by the ubiquitin-proteasome pathway, depending on the phosphorylation state of Thr58 and Ser-62 with phosphorylated Thr-58 recognized by the Fbxw7 E3 ubiquitin ligase^{31,32}.

Oppositely, the ubiquitin-specific proteases can de-ubiquitinate and protect MYC via counteracting Fbw7alpha and Fbw7gamma^{33,34}. MYC phosphorylation, ubiquitination and proteasome degradation has been extensively targeted in studies

aiming to gain control of MYC levels in cancer cells^{7,14}, with several compounds in preclinical development or early clinical trials^{7,35}, including all-trans retinoic acid (ATRA) and

Vismodegib, which are approved by the FDA for the treatment of acute myeloid leukemia and basal-cell carcinoma^{36,37}. In our study, the effect of rLon was not associated with the phosphorylation state of Thr-58 and Ser-62 in cellular models, in contrast to Omomyc^{38,39}. A tempting explanation for the Lon effect would be that the bacterial Lon peptidase with its inherent function as a scavenger of unfolded proteins, is able to access degraded disordered MYC domains, thus circumventing the specifically mammalian phosphorylation checkpoints that would otherwise protect cancerous dysregulated MYC^{14,40,41}. The added contribution of MYC-MAX inhibition by the Lon NTD, predicted by our *in silico* data, would further lower MYC expression levels, especially for L-MYC, in cells and tissues.

Inhibitors targeting deregulated MYC would ideally affect most human cancers, but selective inhibitors have been difficult to find and off target effects of MYC inhibitors are a significant concern. This study showed that bacterial Lon protease effectively enhances the anti-bacterial defense against kidney infections and prevents tissue pathology, by preventing MYC deregulation in infected tissues. Remarkably, rLon treatment of infected mice resulted in a virtual shut down of renal toxicity genes. Lon also affected c-MYC, N-MYC and L-MYC protein levels in infected as well as in tumor bearing mice.

These potent effects are remarkable, especially since there was no evidence that rLon targeted normal MYC expression. The molecular basis of this apparent specificity remains unclear, but several scenarios can be proposed. First, the Lon protease may target disease-specific modifications of MYC that optimize Lon binding and the degradation of its target^{14,40,41}. Second, MYC overexpression may make it more available for the protease to execute its proteolytic effect⁴¹. Third, access of Lon, specifically the NTD and AAA+ domains, to MYC promoters may be enhanced in the open chromatin of rapidly dividing cells

in cancer or infection, possibly accounting for the pronounced transcriptional inhibition detected in diseased tissues but not in healthy tissues⁴²

References

- 1 Dang, C. V. MYC on the path to cancer. *Cell* **149**, 22-35 (2012).
- 5 2 Zeller, K. I. *et al.* Global mapping of c-Myc binding sites and target gene networks in human B cells. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 17834-17839 (2006).
- 3 Wierstra, I. & Alves, J. The c.myc promoter: Still mysterY and challenge. *Adv. Cancer Res.* **99**, 113-333 (2008).
- 4 Pelengaris, S., Khan, M. & Evan, G. c-MYC: more than just a matter of life and
10 death. *Nat. Rev. Cancer* **2**, 764-776 (2002).
- 5 Alexandrova, R. & Podlipnik, Č. in *Oncogenic Viruses* (ed Moulay Mustapha Ennaji) 191-219 (Academic Press, 2023).
- 6 Dong, Y., Tu, R., Liu, H. & Qing, G. Regulation of cancer cell metabolism: oncogenic MYC in the driver's seat. *Signal Transduct. Target Ther.* **5**, 124 (2020).
- 15 7 Whitfield, J. R., Beaulieu, M.-E. & Soucek, L. Strategies to inhibit Myc and their clinical applicability. *Front. Cell Dev. Biol.* **5**, 10 (2017).
- 8 Annibali, D. *et al.* Myc inhibition is effective against glioma and reveals a role for Myc in proficient mitosis. *Nat. Commun.* **5**, 4632 (2014).
- 9 Soucek, L. *et al.* Inhibition of Myc family proteins eradicates KRas-driven lung cancer
20 in mice. *Genes Dev.* **27**, 504-513 (2013).
- 10 Soucek, L., Nasi, S. & Evan, G. Omomyc expression in skin prevents Myc-induced papillomatosis. *Cell Death Differ.* **11**, 1038-1045 (2004).
- 11 Massó-Vallés, D. *et al.* MYC inhibition halts metastatic breast cancer progression by blocking growth, invasion, and seeding. *Cancer Res. Commun.* **2**, 110-130 (2022).
- 25 12 Wang, E. *et al.* Tumor penetrating peptides inhibiting MYC as a potent targeted therapeutic strategy for triple-negative breast cancers. *Oncogene* **38**, 140-150 (2019).
- 13 Duffy, M. J., O'Grady, S., Tang, M. & Crown, J. MYC as a target for cancer treatment. *Cancer Treat. Rev.* **94**, 102154 (2021).

- 14 Llombart, V. & Mansour, M. R. Therapeutic targeting of “undruggable” MYC.
EBioMedicine **75** (2022).
- 15 Lu, B. Mitochondrial Lon protease and cancer. *Mitochondrial DNA and Diseases*, 173-182 (2017).
- 5 16 Lee, I. & Suzuki, C. K. Functional mechanics of the ATP-dependent Lon protease-lessons from endogenous protein and synthetic peptide substrates. *Biochim. Biophys. Acta, Proteins Proteomics* **1784**, 727-735 (2008).
- 17 Venkatesh, S., Lee, J., Singh, K., Lee, I. & Suzuki, C. K. Multitasking in the mitochondrion by the ATP-dependent Lon protease. *Biochim. Biophys. Acta, Mol. Cell Res.*
10 **1823**, 56-66 (2012).
- 18 Butler, D. S. C. *et al.* A bacterial protease depletes c-MYC and increases survival in mouse models of bladder and colon cancer. *Nat. Biotechnol.* **39**, 754764 (2021).
- 19 Mugrauer, G. & Ekblom, P. Contrasting expression patterns of three members of the myc family of protooncogenes in the developing and adult mouse kidney. *J. Cell Biol.* **112**,
15 13-25 (1991).
- 20 Tullus, K. & Shaikh, N. Urinary tract infections in children. *Lancet* **395**, 16591668 (2020).
- 21 Johnson, J. R. & Russo, T. A. Acute pyelonephritis in adults. *New Eng. J. Med.* **378**, 48-59 (2018).
- 20 22 Prasada, S., Bhat, A., Bhat, S., Shenoy Mulki, S. & Tulasidas, S. Changing antibiotic susceptibility pattern in uropathogenic *Escherichia coli* over a period of 5 years in a tertiary care center. *Infect. Drug Resist.*, 1439-1443 (2019).
- 23 Castell, A. *et al.* A selective high affinity MYC-binding compound inhibits MYC: MAX interaction and MYC-dependent tumor cell proliferation. *Sci. Rep.* **8**, 10064 (2018).
- 25 24 Yang, J., Cheng, W.-x., Wu, G., Sheng, S. & Zhang, P. Prediction of folding patterns for intrinsic disordered protein. *Sci. Rep.* **13**, 20343 (2023).
- 25 Karlowicz, A. *et al.* Defining the crucial domain and amino acid residues in bacterial Lon protease for DNA binding and processing of DNA-interacting substrates. *J. Biol. Chem.* **292**, 7507-7518 (2017).
- 30 26 Olivares, A. O., Baker, T. A. & Sauer, R. T. Mechanistic insights into bacterial AAA+ proteases and protein-remodelling machines. *Nat. Rev. Microbiol.* **14**, 3344 (2016).

- 27 Amati, B. *et al.* Oncogenic activity of the c-Myc protein requires dimerization with Max. *Cell* **72**, 233-245 (1993).
- 28 Fischer, H. *et al.* Pathogen specific, IRF3-dependent signaling and innate resistance to human kidney infection. *PLoS Pathog.* **6**, e1001109 (2010).
- 5 29 Puthia, M. *et al.* IRF7 inhibition prevents destructive innate immunity—A target for nonantibiotic therapy of bacterial infections. *Sci. Transl. Med.* **8**, 336ra359336ra359 (2016).
- 30 Roberts, J. A. Management of pyelonephritis and upper urinary tract infections. *Urol. Clin. North Am.* **26**, 753-763 (1999).
- 31 Yada, M. *et al.* Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. *EMBO J.* **23**, 2116-2125 (2004).
- 10 32 Welcker, M. *et al.* The Fbw7 tumor suppressor regulates glycogen synthase kinase 3 phosphorylation-dependent c-Myc protein degradation. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 9085-9090 (2004).
- 33 Popov, N. *et al.* The ubiquitin-specific protease USP28 is required for MYC stability.
- 15 34 Sun, X. X. *et al.* The nucleolar ubiquitin-specific protease USP36 deubiquitinates and stabilizes c-Myc. *Proc. Natl. Acad. Sci. U. S. A.* **112**, 37343739 (2015).
- 35 Sun, X. X., Li, Y., Sears, R. C. & Dai, M. S. Targeting the MYC ubiquitination proteasome degradation pathway for cancer therapy. *Front. Oncol.* **11**, 679445 (2021).
- 20 36 Giuli, M. V. *et al.* Current trends in ATRA delivery for cancer therapy. *Pharmaceutics* **12**, 707 (2020).
- 37 Axelson, M. *et al.* US Food and Drug Administration approval: vismodegib for recurrent, locally advanced, or metastatic basal cell carcinoma. *Clin. Cancer Res.* **19**, 2289-2293 (2013).
- 25 38 Truica, M. I., Burns, M. C., Han, H. & Abdulkadir, S. A. Turning up the heat on MYC: Progress in small-molecule inhibitors. *Cancer Res.* **81**, 248-253 (2021).
- 39 Han, H. *et al.* Small-molecule MYC inhibitors suppress tumor growth and enhance immunotherapy. *Cancer Cell* **36**, 483-497.e415 (2019).

- 40 Madden, S. K., de Araujo, A. D., Gerhardt, M., Fairlie, D. P. & Mason, J. M. Taking
the Myc out of cancer: toward therapeutic strategies to directly inhibit cMyc. *Mol. Cancer*
20, 3 (2021).
- 41 Wolf, E. & Eilers, M. Targeting MYC proteins for tumor therapy. *Annu. Rev. Cancer*
5 *Biol.* **4**, 61-75 (2020).
- 42 Arends, J. *et al.* An integrated proteomic approach uncovers novel substrates and
functions of the Lon protease in *Escherichia coli*. *Proteomics* **18**, 1800080 (2018).
- 43 Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods*.
19, 679-682 (2022).
- 10 44 Mobley, H. *et al.* Pyelonephritogenic *Escherichia coli* and killing of cultured human
renal proximal tubular epithelial cells: role of hemolysin in some strains. *Infect. Immun.* **58**,
1281-1289 (1990).
- 45 Sato, M. *et al.* Distinct and essential roles of transcription factors IRF-3 and IRF-7 in
response to viruses for IFN- α/β gene induction. *Immunity* **13**, 539-548 (2000).
- 15 46 Edén, C. S. *et al.* Inhibition of experimental ascending urinary tract infection by an
epithelial cell-surface receptor analogue. *Nature* **298**, 560-562 (1982).
- 47 Moser, A. R., Pitot, H. C. & Dove, W. F. A dominant mutation that predisposes to
multiple intestinal neoplasia in the mouse. *Science* **247**, 322-324 (1990).
- 48 Su, L.-K. *et al.* Multiple intestinal neoplasia caused by a mutation in the murine
20 homolog of the APC gene. *Science* **256**, 668-670 (1992).

49

CLAIMS

1. A peptide comprising an amino acid sequence having substantial identity to the amino acid sequence of one or more domains of Lon protease or a functional fragment or
5 variant thereof.
2. The peptide of claim 1, wherein the Lon protease domain is the N-terminal domain.
3. The peptide of claim 2, wherein the N-terminal domain comprises part or all of the alpha helices region.
4. The peptide of claim 2, wherein the N-terminal domain omits part or all of the alpha
10 helices region.
5. The peptide of claim 1, wherein the Lon protease domain is the AAA+ domain.
6. The peptide of claim 1, wherein the Lon protease domain is the C-terminal proteolytic domain.
7. The peptide of any preceding claim, wherein the peptide comprises an amino acid
15 sequence having substantial identity to the amino acid sequence of a first Lon protease domain and an amino acid sequence having substantial identity to the amino acid sequence of a second Lon protease domain.
8. The peptide of any preceding claim, wherein the peptide comprises an amino acid sequence having substantial identity to only one domain, or only two domains of Lon
20 protease.
9. The peptide of any preceding claim, wherein the peptide comprises an amino acid sequence having substantial identity to the amino acid sequence of the N-terminal domain of Lon protease and an amino acid sequence having substantial identity to the amino acid sequence of another domain.
- 25 10. The peptide of any preceding claim, wherein the peptide comprises an amino acid sequence having substantial identity to the amino acid sequence of the C-terminal proteolytic of Lon protease and an amino acid sequence having substantial identity to the amino acid sequence of another domain.
- 30 11. The peptide of any preceding claim, wherein the peptide comprises an amino acid sequence having substantial identity to only one domain, or only two domains of Lon protease.

12. The peptide of any preceding claim, wherein the peptide does not comprise the amino acid sequence of whole wild-type Lon protease.
13. The peptide of any preceding claim, wherein the Lon protease is of bacterial origin.
14. The peptide of any preceding claim, wherein the peptide is less than about 780 amino acids in length.
15. The peptide according to any preceding claim, wherein the peptide is more than about 110 amino acids in length.
16. The peptide of any of claims 1 to 4 and 7 to 15, comprising an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID No. 2 or 3 or a fragment or variant thereof.
17. The peptide of any of claims 1, 5 and 7 to 15, comprising an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID No. 4 or a fragment or variant thereof.
18. The peptide of any of claims 1 and 6 to 15 comprising an amino acid sequence having substantial identity to the amino acid sequence of SEQ ID No. 5 or a fragment or variant thereof.
19. The peptide, variant or fragment may be isolated, substantially isolated, purified or substantially purified. A peptide, variant or fragment is isolated or purified if it is completely free of any other components, such as lipids or other peptides. A peptide, variant or fragment is substantially isolated if it is mixed with carriers or diluents which will not interfere with its intended use. For instance, a peptide, variant or fragment is substantially isolated or substantially purified if it is present in a form that comprises less than 10%, less than 5%, less than 2% or less than 1% of other components.
20. A polynucleotide encoding a peptide according to any preceding claim.
21. A vector comprising the polynucleotide of claim 20.
22. A host cell comprising the polynucleotide of claim 20 or the vector of claim 21.
23. A pharmaceutical composition comprising a peptide according to any of claims 1 to 19, or a polynucleotide according to claim 20 or a vector according to claim 21, and a pharmaceutically acceptable carrier.
24. The peptide according to any of claims 1 to 19, the polynucleotide according to claim 20, the vector according to claim 21 or the pharmaceutical composition according to claim 23 for use in therapy.

25. The peptide according to any of claims 1 to 19, the polynucleotide according to claim 20, the vector according to claim 21 or the pharmaceutical composition according to claim 23 for use in the treatment or prevention of infections, especially bacterial infections, or the treatment of cancer or for use in the treatment of a disease modulated by L-MYC, N-MYC or C-MYC, or for use in regulating L-MYC, N-MYC or C-MYC.
26. A method of treating or preventing an infection or cancer, or treating a disease modulated by L-MYC, N-MYC or C-MYC, or regulating L-MYC, N-MYC or C-MYC comprising the step of administering an effective amount of peptide according to any of claims 1 to 19, the polynucleotide according to claim 20, the vector according to claim 21 or the pharmaceutical composition according to claim 23 to a subject in need thereof.
27. A peptide comprising an amino acid sequence havthe amino acid sequence of Lon Protease or a functional variant or fragment thereof for use in the treatment of a disease, such as an infection or cancer, wherein the disease is modulated by L-MYC or N-MYC.