Supporting Information

Perera et al. 10.1073/pnas.1200470109

SI Materials and Methods

Preparation of NSP4 Inclusion Body Material. The DNA sequence encoding human neutrophil serine protease 4 (NSP4) was amplified from human bone marrow cDNA using primers DJ2064B 5′-GGCAGATCATATGGAAATCATCGGCGGCCACGAAGTG-ACC-3′ and DJ1715 5′-GGGAATTCTTATTAACGACGAAC-CACGTCCCA-3′ (Metabion). The PCR product was digested with *Nde* II/*Eco* RI and subcloned into the expression vector pET24c(+) (Novagen). Positive clones were verified by sequence analysis and transformed into the *Escherichia coli* strain B834 (DE3) (Novagen).

Overnight cultures of transformed *E. coli* strain B834 were grown in Luria-Bertani (LB) broth containing 50 μg/mL kanamycin, 2% (wt/vol) glucose, and inoculated in LB/kanamycin without glucose. Expression was induced at an $A_{600\,nm}$ of 0.5–0.8 by adding isopropyl-1-thio-β-D-galactopyranoside to a final concentration of 1 mM. Cultures were grown shaking at 150 rotations per minute at 37 °C for 3–4 h after induction and harvested by centrifugation.

Pelleted bacteria were lysed in 50 mM Tris-HCl, 2 mM $MgCl_2$ containing 10 μg/mL DNase I and 0.25 mg/mL lysozyme, pH 7.2, by sonification. Inclusion body (IB) material was obtained by centrifugation and two washing steps in 50 mM Tris-HCl, 60 mM EDTA, 1.5 M NaCl, 6% Triton X-100, pH 7.2, and in 50 mm Tris-HCl, 60 mm EDTA, pH 7.2.

Solubilization of IB was performed in an end-over-end rotator in 6 M guanidinium hydrochloride, 100 mM Tris-HCl, 20 mM EDTA, 15 mM GSH, 150 mM GSSG, pH 8.0, overnight at room temperature. Solubilized IB were then dialized against 6 M guanidinium hydrochloride, pH 5.0, at 4 °C. Protein concentration was determined by measuring the absorbance at 280 nm (spectrophotometer ND-1000; Peqlab).

Expression of S-Tag-NSP4 in HEK 293 Cells Using the "Flp-In" System. The pcDNA5/FRT expression vector (Invitrogen) was complemented by an Igκ-chain secretion signal using the oligoduplex DJ2972 5′-Pho-CTAGCCACCATGGAGACAGACACACTCC-TGCTATGGGTACTGCTGCTCTGGGTACCAGGTTCCAC-3′ and DJ2973 5′-Pho-GTGGAACCTGGTACCCAGAGCAGCA-GTACCCATAGCAGGAGTGTGTCTGTCTCCATGGTGG-3′ (Metabion), followed by an S-tag peptide using the oligoduplex DJ3151 5′-Pho-GTGAAAGAAACCGCTGCTGCTAAATTCG-AACGCCAGCACATGGACAGCGGAAGCGGTGGACAC-3′ and DJ3152 5′-Pho-GTGTCCACCGCTTCCGCTGTCCATGT-GCTGGCGTTCGAATTTAGCAGCAGCGGTTTCTTTCAC-3′. The cDNA-sequence of mature NSP4, starting at the conserved Ile-Ile-Gly-Gly-motif, was amplified from human bone marrow cDNA with DJ3215 5′-TGGACACGTGGATGACGACGACAAGATC-ATCGGGGGCCACGAG-3′ and DJ3216 5′-CCCCACCGGTC-TTCCGAACCACGTCCCAGATCCA-3′ (Metabion), thereby introducing an enterokinase cleavage site (DDDDK) at the mature amino terminus. The resulting PCR product was digested with *Pml* I/*Age* I and cloned into the reading frames of the N-terminal S-tag and the C-terminal 6xHis-tag of the pcDNA5/FRT vector. The construct was verified by sequence analysis.

HEK 293 cells (Invitrogen) were cultured in DMEM (Life Technologies) supplemented with 10% FCS. One day before transfection, 1.4×10^6 flp-in HEK 293 cells were plated in 35-mm cell culture dishes and incubated in 3 mL DMEM containing 10% FCS. Cells were transfected by adding 1.8 mg pOG44 (encoding the flp-in recombinase; Invitrogen) and 0.2 mg pcDNA5/FRT/S-tag-NSP4/H6 in 100 mL serum-free OptiMEM medium and 7 mL FuGENE HD transfection reagent (Roche). Starting at 48 h after transfection, the medium was replaced with DMEM containing 10% FCS and 75 μg/mL hygromycin B (Invitrogen) every 2–3 d. Two weeks posttransfection, visible circular hygromycin B-resistant colonies were pooled and cultured until a desired number of plates had reached confluence. Expression of recombinant S-tag-NSP4 was tested by Western blot analysis with S-protein conjugated to HRP (Novagen). For protein expression, cells were cultured in DMEM supplemented with 5% FCS and 75 μg/mL hygromycin B for 8–10 d before the supernatant was harvested.

Purification of S-Tag-NSP4 from HEK 293 Supernatant. Harvested cell-culture supernatant was filtered through a 0.22-μm membrane (Millipore), concentrated fivefold, and dialyzed against binding buffer (20 mM Na_2HPO_4; 500 mM NaCl; 50 mM imidazole, pH 7.5) at 4 °C. The 6xHis-tagged S-tag-NSP4 was purified using nickel-affinity chromatography. The protein solution was applied to a HisTrap HP column (GE Healthcare) previously equilibrated in binding buffer. The column was washed with binding buffer, and bound proteins were eluted applying a linear imidazole gradient from 50 mM to 1 M imidazole in 20 mM Na_2HPO_4, 500 mM NaCl, pH 7.5. Fractions were collected and analyzed for S-tag-NSP4 by SDS/PAGE and silver-staining. The expected size is 31 kDa. S-tag-NSP4 containing fractions were pooled and concentrated. Protein concentration was determined by measuring the absorbance at 280 nm.

Processing of S-Tag-NSP4 by Enterokinase. Purified S-tag-NSP4 was dialyzed against 20 mM Tris-HCl, 500 mM NaCl, pH 7.4 at 4 °C. After dialysis, $CaCl_2$ was added to a final concentration of 2 mM. The amino-terminal extension, the S-tag peptide and the enterokinase (EK) recognition sequence, was cleaved off by Tag-Off High Activity Enterokinase (Merck) to generate mature 6xHis-tagged NSP4. EK cleavage was accomplished by 1 U EK per 50 μg S-tag-NSP4 for 16 h at room temperature. Conversion of S-tag-NSP4 was verified by SDS/PAGE and subsequent silverstaining. There is a molecular weight shift of ∼3 kDa after removal of the synthetic propeptide. Converted NSP4 was subsequently purified by nickel-affinity chromatography according to the manufacturers' instructions (Ni-NTA; Qiagen).

Production of Granzymes A, B, K, M, H, and Proteinase 3. Granzymes (Gzms) A, B, K, M, and H were expressed as inclusion body proteins, refolded and converted into the active mature form as previously reported (1–5). Proteinase 3 (PR3) was produced as described by Korkmaz et al. (6).

Production of NSP4 Precursor in HEK 293 EBNA Cells. The DNA sequence encoding the full length reading frame of NSP4 (natural NSP4) was amplified from human bone marrow cDNA using primers DJ539 5′-TCGGCGGCCGCACCACGTCCCAGATC-CAGG-3′ and DJ583 5′-CGGAATTCTGCCATGGGGCTCG-GGTTGA-3′ (Metabion). The PCR product was digested with *Eco* RI and the target vector pcDNA3.1 (Invitrogen) was digested with *Bsp* 120I, complemented by the Klenow enzyme and subsequently digested with *Eco* RI. After subcloning in frame with the C-terminal cMyc-6xHis-tag, positive clones were verified by sequence analysis.

The HEK 293 cell line stably expressing Epstein–Barr virus nuclear antigen-1 (HEK293E) was grown as suspension culture in Freestyle 293 Expression Medium (Invitrogen), supplemented with 0.1% Pluronic F-68, 25 μg/mL Geneticin G418 and 0.5% Bacto TC Lactalbumin Hydrolysate. Before transfection of HEK293E with natural NSP4, cells were brought to 1×10^6 cells/mL. Poly-

ethylene imine (PEI)/DNA complexes were prepared by adding PEI to DNA, both prediluted in Optipro serum free medium (Invitrogen) and incubation for 15 min at room temperature before adding to cell culture. Per 1 mL cell culture, 2 μg PEI and 1 μg DNA were used.

Cell culture supernatant was harvested after 72–96 h by 10-min centrifugation at 1,000 × *g* and purified as described above for S-tag-NSP4.

ELISA. The specificity of anti-NSP4 positive mAbs was tested in ELISA by coating 1 μg NSP4 IB material or 100 ng per well of each of the Gzms A, B, K, M, H, neutrophil elastase (NE), cathepsin G (CG), PR3, azurocidin (AZU), and native NSP4 produced in HEK 293. NE, AZU, and CG were purchased from Elastin products, Sigma-Aldrich, and Calbiochem, respectively. Hybridoma culture supernatants were diluted 1:20 and bound mAbs were detected with HRP-labeled goat anti-rat IgG+IgM antibodies diluted 1:4,000 (112-035-068; Jackson Immunoresearch).

Isolation of Different Blood Cell Populations and Preparation of Total Cell Lysates. Peripheral blood mononuclear cells (PBMCs) and polymorphonuclear cells (PMNs) were obtained by Ficoll gradient centrifugation. Residual erythrocytes were removed by dextran sedimentation and hypertonic treatment. For total cell lysates, cells were taken in 50 mM Tris/HCl, 150 mM NaCl, 0.5 mM EDTA, 1 mM PMSF, 0.5% Nonidet P-40, pH 7.4, incubated on ice for 20 min, and underwent 4 × 30 s ultrasonication (Sonorex TK52; Bandelin Electronic). Cell debris and DNA were pelleted by centrifugation for 10 min at 20,000 × *g* and 4 °C.

SDS/PAGE and Immunoblotting. To examine the specificity of the developed rat mAbs in immunoblotting, 100 ng of each tested protein was separated by reducing SDS/PAGE and analyzed by silver-staining and immunoblotting. For the latter, proteins were transferred to a Hybond-ECL membrane (GE Healthcare) and probed with the mAbs against NSP4 (1:30) and HRP-labeled goat anti-rat antibodies (Jackson Immunoresearch) in a dilution of 1:10,000.

Total cell lysates of isolated PBMCs or PMNs were resolved by reducing SDS/PAGE and analyzed by Western blotting, as described above, using anti-NSP4 mAbs.

Immunohistochemistry. Formalin-fixed, paraffin-embedded human tissue samples from bone marrow, lymph node, spleen, neural network, pancreas, prostate, and arteries were routinely embedded for light microscopy and sectioned into 2-μm slides. Immunostaining was carried out after antigen retrieval by microwaving in citrate buffer (3 × 5 min, 600 W) after dewaxing of the sections. The rat hybridoma supernatants were diluted 1:30 and incubated with the sections overnight at 4 °C. Negative control slides were stained in parallel omitting the incubation with primary antibodies. Slides were developed using the Ultravision LP staining kit (Thermo Fisher Scientific). Sections were counterstained with hematoxylin.

Conversion of NSP4 Precursor by Dipeptidyl Peptidase I. Dipeptidyl peptidase I (DPPI) was first activated at 37 °C in 50 mM sodium acetate, 100 mM NaCl, 1 mM dithioerythritol (DTE), pH 5.5 for 60 min. NSP4 precursor was incubated with 0.5 U/mg activated DPPI in 50 mM sodium acetate, 100 mM NaCl, pH 5.5 for 4 h at room temperature.

Edman Sequencing. Ten micrograms of NSP4 precursor before and after incubation with DPPI was separated by reducing SDS/PAGE and transferred to Immobilon-P transfer membrane (Millipore) in 90 mM Tris-borate, 1 mM EDTA, and 10% methanol. Membrane was stained with Coomassie blue and the bands representing NSP4 were analyzed by N-terminal Edman sequencing according to standard procedure.

1. Andrade F, Fellows E, Jenne DE, Rosen A, Young CS (2007) Granzyme H destroys the function of critical adenoviral proteins required for viral DNA replication and granzyme B inhibition. *EMBO J* 26:2148–2157.
2. Fellows E, Gil-Parrado S, Jenne DE, Kurschus FC (2007) Natural killer cell-derived human granzyme H induces an alternative, caspase-independent cell-death program. *Blood* 110:544–552.
3. Hink-Schauer C, Estébanez-Perpiñá E, Kurschus FC, Bode W, Jenne DE (2003) Crystal structure of the apoptosis-inducing human granzyme A dimer. *Nat Struct Biol* 10:535–540.
4. Kurschus FC, et al. (2004) Killing of target cells by redirected granzyme B in the absence of perforin. *FEBS Lett* 562:87–92.
5. Wilharm E, et al. (1999) Generation of catalytically active granzyme K from *Escherichia coli* inclusion bodies and identification of efficient granzyme K inhibitors in human plasma. *J Biol Chem* 274:27331–27337.
6. Korkmaz B, Kuhl A, Bayat B, Santoso S, Jenne DE (2008) A hydrophobic patch on proteinase 3, the target of autoantibodies in Wegener granulomatosis, mediates membrane binding via NB1 receptors. *J Biol Chem* 283:35976–35982.

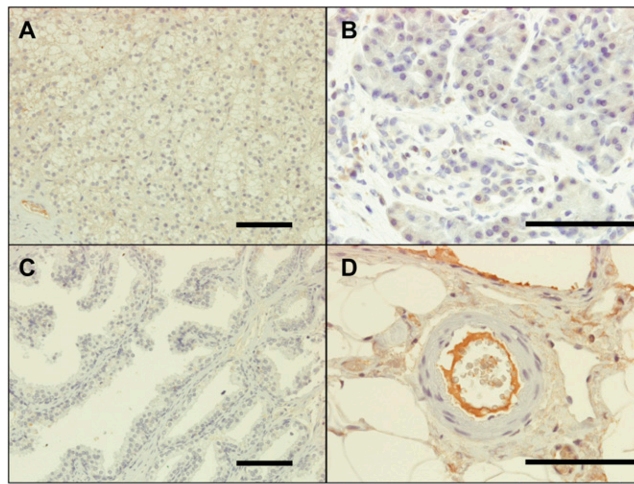

Fig. S1. NSP4 protein expression screen. Staining of different human tissue samples using anti-NSP4 mAbs and the Ultravision LP staining kit. NSP4 was not detected in (*A*) neural network, (*B*) pancreas, (*C*) prostate, or (*D*) arteries. (Scale bars, 100 μm.)

Table S1. NSP4 cleavage sites from a chymotryptic peptide library

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
QFDAAFVPEV	AAQKAPASP	0.86	25.6	−0.7	955.5	17888	
YHGRVQALAD	AAREAGLQF	0.96	25.2	0.6	1049.5	17897	
AHEAGEFFMR	AGSATVRPTEGAGGTL	1.00	45.5	−0.3	1531.7	17875	
AHEAGEFFMR	AGSATVRPTEGAGGTL	1.00	45.6	−1.5	1531.7	17875	
GAVIAYEPVW	AIGTGKSATPAQAQAVH	1.00	52.6	−0.8	1722.9	17903	
GAVIAYEPVW	AIGTGKSATPAQAQAVH	1.00	49.6	−1.7	1722.9	17903	
NFNVPFVVVR	AISDVADQQSHL	1.00	50.3	−0.9	1370.6	17863	
NFNVPFVVVR	AISDVADQQSHL	1.00	53.9	−1.3	1370.6	17863	
MSFELPALPY	AKDALAPH	0.97	32.1	−0.7	937.5	17879	X
MSFELPALPY	AKDALAPH	0.99	29.3	−2.2	937.5	17879	X
KLDLERTVIR	APADGWVTNL	0.84	26.8	2.0	1130.5	17896	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	45.9	2.2	1745.9	17899	
LCRNCKIVKR	DGVIRVIC	0.85	31.6	−0.9	1018.5	17896	
SLXGTIIVGR	DIAHAKL	0.96	36.5	−0.5	882.5	17878	
SLXGTIIVGR	DIAHAKL	0.83	34.3	−0.2	882.5	17878	
KLTKRMRVIR	EKVDATKQY	1.00	35.7	−0.9	1224.6	17904	
KLTKRMRVIR	EKVDATKQY	0.99	32.8	−1.1	1224.6	17904	
NNPFFVSLKD	GAQKEADKLGY	0.95	38.2	−0.3	1322.7	17901	
NNPFFVSLKD	GAQKEADKLGY	0.97	38.1	−0.5	1322.7	17901	
GIANTFIAAQ	GHDVGKSLY	0.82	29.7	1.7	1090.5	17892	
KPEDAVLDVQ	GIATVTPAIVQ	0.97	41.8	−1.2	1156.6	17899	
KPEDAVLDVQ	GIATVTPAIVQ	0.97	38.1	−0.9	1156.6	17899	
KPEDAVLDVQ	GIATVTPAIVQAC	0.90	35	−3.0	1387.7	17899	
GQNVEFEIQD	GQKGPAAVNVTAI	0.90	27.1	−2.2	1340.7	17881	
HVIAGKAVAL	KEAMEPEFKTY	0.85	36.9	−2.1	1515.7	17889	X
SGVRAIDTKC	KIEQAPGQH	0.99	28.1	−0.9	1122.5	17896	
PDEGIPAVCF	KLKDGEDPGY	0.99	27.9	−0.8	1264.6	17877	X
PDEGIPAVCF	KLKDGEDPGY	0.94	31.1	−2.7	1264.6	17877	X
PDEGIPAVCF	KLKDGEDPGY	0.97	30.8	−2.6	1264.6	17877	X
KNPQKNLYTF	KNQASNDLPN	1.00	24.6	−1.1	1215.6	87082	X
KNPQKNLYTF	KNQASNDLPN	0.97	19.9	−1.1	1215.6	87082	X
SGGAGIQADL	KTFSALGAY	0.81	30.9	−7.4	1072.5	17884	X
MVEVFLERGY	KVVSGGTDNHLF	0.99	41.4	−0.6	1388.7	17889	X
AKSVKFKYPR	QRKTVVADGVGQGY	0.93	29.2	−3.0	1592.8	17863	
WRQKGTGRAR	SGSIKSPIW	0.89	30.7	0.0	1089.6	17897	
ARPPLVVISR	SHADAELKEY	1.00	42.2	−2.0	1277.6	17906	
ARPPLVVISR	SHADAELKEY	1.00	42.1	−2.0	1277.6	17906	
KEASAGKLVR	TLAAVRDAKEAA	1.00	42.3	−1.9	1330.7	17904	
KEASAGKLVR	TLAAVRDAKEAA	1.00	39.5	−1.0	1330.7	17904	
LSQYDFPGDD	TPIVRGSAL	0.97	31	−1.0	1000.5	17897	
RDXKGAVASL	TSVAKLPF	0.83	24.9	0.5	977.5	17871	X
VVNTIRGIVK	VAAVKAPGF	1.00	41.8	−1.2	974.5	17905	
KHLPEPFRIR	VIEPVKR	0.91	41.5	1.3	955.6	87082	
KHLPEPFRIR	VIEPVKR	0.98	41.7	−0.7	955.6	87082	
DIVWDFRLPR	VKSISASGH	0.98	33.5	−1.4	1000.5	17877	
MSTAKL	VKSKATNLLY	0.99	27.3	−0.6	1279.7	17870	X
CKPTSPGRRH	VVKVVNPELH	0.96	29.1	−1.9	1248.7	17897	
CKPTSPGRRH	VVKVVNPELH	0.94	29.4	−1.0	1248.7	17897	

Nonprime-side sequences are reported up to the position P10. X denotes ambiguous nonprime side residues. A small number of cleavage sites have P1 residues, which correspond to a potential cleavage-site of the digestion protease used for library generation. These cases might be related to incomplete amine protection during library preparation. Hence, prime-side cleavage products with such amino-termini are omitted from the proteomic identification of protease cleavage sites (PICS) analysis, as well as prime-side cleavage products with amino termini that correspond to a protein amino terminus. One exemplary protein is stated for each identified peptide sequence. X!Tandem was applied to assign peptides to tandem mass spectra and the X!Tandem hyperscore is reported together with the PeptideProphet probability score. C signifies carboxyamidomethylated cysteine and K signifies di-methylated lysine. Some peptides were identified repeatedly by LC-MS/MS; however, the set of sequences was rendered nonredundant before analysis in the form of sequence logos.

Table S2. NSP4 cleavage sites from a GluC peptide library

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
GYTLNGRTIR	AAMVTVAKAKA	1.00	43.6	−2.3	1203.7	17889	
GYTLNGRTIR	AAMVTVAKAKA	1.00	50.5	−1.6	1203.7	17889	
KYYDPRVWLR	AGQTSMIARLE	0.93	37.1	−0.6	1263.6	17892	
KYYDPRVWLR	AGQTSMIARLE	1.00	48.2	−0.2	1263.6	17892	
KYYDPRVWLR	AGQTSMIARLE	0.97	40.8	−0.6	1263.6	17892	
SQASDSYYYR	AKQAVYR	0.78	25.4	−1.4	950.5	17886	
PFLEHDDANR	ALMGANMQR	0.99	38	−0.7	1078.5	17904	
PFLEHDDANR	ALMGANMQR	1.00	41.9	−2.1	1078.5	17904	
FQPTAVGFAE	ALNNKDKPE	0.80	37.3	−0.6	1171.6	17899	X
KLEHAVPMAK	ALVAGGVRVLE	1.00	38	−2.8	1170.6	17881	
KLEHAVPMAK	ALVAGGVRVLE	1.00	49.3	−0.4	1170.6	17881	
APDQTTTIVR	ANSSTTTAAEPLKM	0.93	30.3	−0.8	1536.7	17886	
APDQTTTIVR	ANSSTTTAAEPLKM	0.96	37	−1.0	1536.7	17886	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	53.5	0.0	1745.9	17899	
HVSPGALDAE	AYGVKSTIED	1.00	35.1	7.3	1197.6	17905	X
GCVELVQPGG	FDELVQIYE	0.91	27.5	0.5	1242.5	14569	
VALKEAMEPE	FKTYQQQVAK	1.00	33.1	0.4	1383.7	17889	X
LVVSTLNNPF	FVSLKDGAQKE	1.00	46.1	−4.4	1364.7	17901	
LVVSTLNNPF	FVSLKDGAQKE	1.00	42.3	−3.6	1364.7	17901	
LTASTVIQYR	GAIIPREAMPA	0.96	29.6	−0.5	1212.6	87082	
ESGVLNQQPY	GFNTRFE	0.76	23.4	−2.2	957.4	17877	
ESGVLNQQPY	GFNTRFE	1.00	27.3	0.3	957.4	17877	
GNGDCHIILR	GGKEPNYSAK	0.90	30.1	−1.5	1193.6	17869	
GNGDCHIILR	GGKEPNYSAK	0.97	33.3	0.3	1193.6	17869	
TTHKTLAGPR	GGLILAKGGSEE	1.00	37.2	−0.4	1245.6	17889	
TTHKTLAGPR	GGLILAKGGSEE	1.00	37.4	−0.5	1245.6	17889	
TEEYGNMIDM	GILDPTKVTR	0.97	43.4	−1.5	1214.7	17905	
TEEYGNMIDM	GILDPTKVTR	0.82	36.2	−2.5	1214.7	17905	
GAILVVAATD	GPMPQTRE	0.99	34.8	0.3	1002.4	17897	
YTSPRVMQAQ	GSQLTNKYAE	0.93	29.5	−0.9	1225.6	17889	
IEDVFSISGR	GTVVTGRVE	0.97	38	−1.3	1004.5	17897	
IEDVFSISGR	GTVVTGRVE	0.96	41.7	−2.4	1004.5	17897	
ARGAIFGLTR	GVNANHIIR	0.82	38.9	−2.3	1080.5	17903	
ARGAIFGLTR	GVNANHIIR	1.00	50.6	−1.4	1080.5	17903	
GDGVGMAIRA	GVPVQDM	0.80	22.6	−1.6	832.3	17869	
QLQQVLMMSR	HNLRAPLANNGSVLE	0.75	37.4	1.9	1691.8	17872	
QLQQVLMMSR	HNLRAPLANNGSVLE	1.00	52.6	1.0	1691.8	17872	
SNYFFDTTQG	HSQINGCTVR	0.77	25.4	−0.8	1258.6	87082	
IVGGGIANTF	IAAQGHDVGKSLYE	0.99	38.6	0.4	1602.8	17892	
IVGGGIANTF	IAAQGHDVGKSLYE	1.00	38.8	−0.1	1602.8	17892	
WWMLHEETVY	KGGDTVTLNE	0.75	28.5	−0.9	1148.5	87082	
GKSASAKSLF	KLQTLGLTQ	1.00	35.2	−1.1	1116.6	17887	
GKSASAKSLF	KLQTLGLTQ	1.00	31.7	−0.8	1116.6	17887	
KNPQKNLYTF	KNQASNDLPN	1.00	33.4	−1.6	1215.6	87082	
KNPQKNLYTF	KNQASNDLPN	0.99	21.7	−2.3	1215.6	87082	
KNPQKNLYTF	KNQASNDLPN	1.00	30	−1.4	1215.6	87082	
VYIEGQLRTR	KWTDQSGQDRYTTE	1.00	38.5	−0.3	1829.8	17904	
VYIEGQLRTR	KWTDQSGQDRYTTE	1.00	42	0.4	1829.8	17904	
ETGYSNKVLD	LIAHISK	0.91	34.5	−0.8	896.5	17880	
QKEADKLGYN	LVVLDSQNNPAKE	1.00	37.5	−0.7	1541.8	17901	
SNYQPSPMVR	MMKADAAPVSAQE	0.99	35.2	1.5	1463.6	17892	
SNYQPSPMVR	MMKADAAPVSAQE	1.00	35	0.0	1463.6	17892	
AQKEADKLGY	NLVVLDSQNNPAKE	0.95	37.6	0.9	1655.8	17901	
AQKEADKLGY	NLVVLDSQNNPAKE	1.00	37.9	−0.2	1655.8	17901	
AQKEADKLGY	NLVVLDSQNNPAKE	1.00	44.7	−0.2	1655.8	17901	
KDPLDNTYTR	NMYIVKH	0.99	32	−0.6	1019.5	17874	
KDPLDNTYTR	NMYIVKH	1.00	28.4	−0.1	1019.5	17874	
AESVKNIFGY	QYTIPTHQGRGAE	0.95	35.1	−0.3	1544.7	87082	
AESVKNIFGY	QYTIPTHQGRGAE	1.00	39.2	−0.9	1544.7	87082	
GINCNLTLLF	SFAQARACAE	1.00	31.2	−2.7	1197.5	17861	

Table S2. Cont.

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
YAKDGISYTF	SIVPNALGKDDE	0.91	36.7	−2.3	1372.7	17871	
YAKDGISYTF	SIVPNALGKDDE	0.99	40.1	0.6	1372.7	17871	
EMFGYATQLR	SLTKGRASYTME	0.78	32.3	−0.1	1458.7	17897	
KGLNVMQNLL	TAHPDVQAVFAQNDE	1.00	34.4	4.9	1728.8	17901	
KGLNVMQNLL	TAHPDVQAVFAQNDE	1.00	48.3	−2.1	1728.7	17901	
KGLNVMQNLL	TAHPDVQAVFAQNDE	1.00	41	−2.0	1728.7	17901	
CKKNPQKNLY	TFKNQASNDLPN	0.91	24.4	−0.7	1463.7	87082	
CKKNPQKNLY	TFKNQASNDLPN	1.00	30.8	1.3	1463.7	87082	
CKKNPQKNLY	TFKNQASNDLPN	0.98	24	−0.3	1463.7	87082	
SVKNIFGYQY	TIPTHQGR	0.93	33.5	−1.2	996.5	87082	
SVKNIFGYQY	TIPTHQGR	1.00	37.3	−0.3	996.5	87082	
SVKNIFGYQY	TIPTHQGRGAE	0.76	34.3	−1.4	1253.6	87082	
KEASAGKLVR	TLAAVRDAKEAA	0.95	34.8	−1.4	1330.7	17904	
VQVNEADALY	TNPAQARE	1.00	41.3	−2.3	973.4	87082	
VQVNEADALY	TNPAQARE	1.00	38.6	−2.2	973.4	87082	
EHAVPMAKAL	VAGGVRVLE	1.00	47.4	−1.0	986.5	17881	
MVMPGDNIKM	VVTLIHPI	1.00	28.6	3.2	978.6	17897	
GVKLINAVQD	VYLDSKITIA	0.91	25.8	−1.6	1237.7	87082	
KTTTTERILF	YTGVNHKIGE	0.96	34.8	−0.1	1232.6	17897	
KTTTTERILF	YTGVNHKIGE	0.95	32.9	0.2	1232.6	17897	

Nonprime-side sequences are reported up to the position P10. X denotes ambiguous nonprime side residues. A small number of cleavage sites have P1 residues, which correspond to a potential cleavage-site of the digestion protease used for library generation. These cases might be related to incomplete amine protection during library preparation. Hence, prime-side cleavage products with such amino-termini are omitted from the PICS analysis, as well as prime-side cleavage products with amino termini that correspond to a protein amino terminus. One exemplary protein is stated for each identified peptide sequence. X!Tandem was applied to assign peptides to tandem mass spectra and the X!Tandem hyperscore is reported together with the PeptideProphet probability score. C signifies carboxyamidomethylated cysteine and K signifies di-methylated lysine. Some peptides were identified repeatedly by LC-MS/MS; however, the set of sequences was rendered nonredundant before analysis in the form of sequence logos.

Table S3. NSP4 cleavage sites from a tryptic peptide library

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
VHLEGGFVGM	AAAPSGASTGSR	1.00	58.1	1.0	1119.5	17891	
VHLEGGFVGM	AAAPSGASTGSR	1.00	55	−0.5	1119.5	17891	
EEAEKITTVQ	AAIDYINGHQA	1.00	41.7	0.5	1259.6	17873	
EEAEKITTVQ	AAIDYINGHQA	1.00	45.3	−1.5	1259.6	17873	
HLEGGFVGMA	AAPSGASTGSR	0.98	33	−0.5	1048.5	17891	
LEEIGVVCQQ	AGGHAAFVDAGK	1.00	57.2	−2.8	1215.6	87082	
LEEIGVVCQQ	AGGHAAFVDAGK	1.00	53.8	−0.5	1215.6	87082	
EWAKHNINVN	AIAPGYMATNNTQQLR	1.00	77.6	0.4	1835.9	17892	
EWAKHNINVN	AIAPGYMATNNTQQLR	1.00	49.6	−3.1	1835.9	17892	
VDHGKTTLTA	AITTVLAK	0.65	33	−1.4	931.5	17897	
VDHGKTTLTA	AITTVLAK	0.85	32.8	−7.5	931.5	17897	
HIPADQFPAQ	ALACELYK	0.74	39.1	−1.9	1082.5	87082	
HIPADQFPAQ	ALACELYK	0.99	37.9	−0.4	1082.5	87082	
HIPADQFPAQ	ALACELYK	1.00	41.8	0.2	1082.5	87082	
EVMGGLGGFG	ALCALPQK	0.81	30.8	−3.3	1015.5	17888	
GDGTTTATVL	AQAIITEGLK	0.90	45.5	−0.1	1158.6	17905	
GDGTTTATVL	AQAIITEGLK	0.83	49.5	4.0	1158.6	17905	
LAVDESFQPT	AVGFAEALNNK	1.00	49.6	2.3	1248.6	17899	
LAVDESFQPT	AVGFAEALNNK	0.99	45.4	0.6	1248.6	17899	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	53.4	0.4	1745.9	17899	
LAVDESFQPT	AVGFAEALNNKDKPE	1.00	57.1	0.4	1745.9	17899	
NPDEAVAIGA	AVQGGVLTGDVK	1.00	61.9	−1.8	1258.7	17861	
DLNPKAMTPV	AWWMLHEE	1.00	38	0.0	1188.5	87082	
KVQNASYQVA	AYLADEIAK	1.00	41.7	−1.1	1108.5	17877	
KVQNASYQVA	AYLADEIAK	0.84	35.2	0.1	1108.5	17877	
GTNTIGSSEA	CMLGGMAMK	1.00	44.8	0.4	1113.5	17877	
GTNTIGSSEA	CMLGGMAMK	1.00	37.9	−3.0	1113.5	17877	
TPTRHYAHVD	CPGHADYVK	1.00	42.4	−4.3	1161.5	17897	
TPTRHYAHVD	CPGHADYVK	1.00	42.5	−1.7	1161.5	17897	
GNARQNLATF	CQTWDDENVHK	1.00	45.2	−3.1	1546.6	17877	
GNARQNLATF	CQTWDDENVHK	1.00	51.6	−1.4	1546.6	17877	
DYDRITKLAR	EAVEGAKL	0.99	31	−0.5	931.5	17881	X
DYDRITKLAR	EAVEGAKL	1.00	37.8	−2.1	931.5	17881	X
ESKKGYINSL	GALTGGQALQQAK	1.00	65.8	−2.1	1357.7	17904	
ESKKGYINSL	GALTGGQALQQAK	1.00	65.8	−1.6	1357.7	17904	
VDESFQPTAV	GFAEALNNKDKPE	0.97	41.7	2.7	1575.8	17899	
VDESFQPTAV	GFAEALNNKDKPE	0.87	45.9	0.9	1575.8	17899	
APAELLFEEF	GFTVDNVVAK	1.00	38	−0.9	1164.6	48994	
APAELLFEEF	GFTVDNVVAK	1.00	41.1	−1.6	1164.6	48994	
KTDQKVVTLS	GFVESQAQAEEAVK	1.00	49.6	−0.8	1607.7	17908	
KTDQKVVTLS	GFVESQAQAEEAVK	1.00	57.3	2.7	1607.7	17908	
RGWQVPAFTL	GGEATDIVVMR	0.99	38	0.6	1234.6	17877	
RGWQVPAFTL	GGEATDIVVMR	1.00	41.9	−0.4	1234.6	17877	
EAVAIGAAVQ	GGVLTGDVK	0.90	38.3	−1.1	960.5	17861	
KYLSDHPKLQ	GIAQQNSFK	0.88	35.2	4.7	1107.5	17877	
KPEDAVLDVQ	GIATVTPAIVQACTQDK	1.00	45.7	−2.5	1888.0	17899	
KPEDAVLDVQ	GIATVTPAIVQACTQDK	1.00	57.1	−0.7	1887.9	17899	
KPEDAVLDVQ	GIATVTPAIVQACTQDK	1.00	53.3	−1.6	1887.9	17899	
VGVDVVAEAT	GLFLTDETAR	0.78	29.1	−1.7	1209.6	17880	
VGVDVVAEAT	GLFLTDETAR	1.00	42.7	3.7	1209.6	17880	
YLFVDMAHVA	GLVAAGVYPNPVPH	1.00	54.2	3.4	1477.7	17889	
YLFVDMAHVA	GLVAAGVYPNPVPH	1.00	50.1	0.0	1477.7	17889	
ITEGLKAVAA	GMNPMDLK	0.98	36.5	1.3	1020.4	17905	
ITEGLKAVAA	GMNPMDLK	0.95	36.5	−2.6	1020.4	17905	
CIAAGIASLW	GPAHGGANEAALK	1.00	72	−1.8	1307.6	17869	
CIAAGIASLW	GPAHGGANEAALK	1.00	75.9	−0.9	1307.6	17869	
NAAADLAAIS	GQKPLITK	0.77	23.1	−3.5	1027.6	17897	
KLGPYEFICT	GRPDEGIPAVCFK	0.83	36.6	−2.5	1560.7	17877	
KLGPYEFICT	GRPDEGIPAVCFK	1.00	39.7	1.1	1560.7	17877	
NYRNHFVTIL	GTIQGEQPGFINK	1.00	58.9	−0.3	1503.7	87082	
NYRNHFVTIL	GTIQGEQPGFINK	1.00	53.3	0.4	1503.7	87082	
FRNAEFLQAY	GVAIADGPLK	0.82	38.6	−0.2	1055.6	17875	

Table S3. Cont.

Nonprime sequence (inferred from database)	Prime-side sequence (LC-MS/MS)	PeptideProphet probability score	X!Tandem hyperscore	Mass accuracy (ppm)	Neutral mass (Da)	Examplary protein accession	Discarded from PICS profile
KARGITINTS	HVEYDTPTR	0.81	33.2	−2.9	1204.5	17897	
KARGITINTS	HVEYDTPTR	0.97	33.2	−1.6	1204.5	17897	
WDFRLPRVKS	ISASGHK	0.65	23.5	−1.6	814.4	17877	
WDFRLPRVKS	ISASGHK	0.70	26.8	−1.6	814.4	17877	
KNPQKNLYTF	KNQASNDLPN	1.00	26.9	−1.6	1215.6	87082	
KNPQKNLYTF	KNQASNDLPN	1.00	30.1	−2.0	1215.6	87082	
GEIEFWFAMI	KVTTIIVM	0.73	20	3.1	1019.6	48994	
NLKAMYSIAK	KYDIPVVMDSAR	1.00	38.9	1.7	1508.7	87082	X
NLKAMYSIAK	KYDIPVVMDSAR	0.97	33.7	−1.4	1508.7	87082	X
GETEDATIAD	LAVGTAAGQIK	0.89	37.1	−0.1	1143.6	17891	
GETEDATIAD	LAVGTAAGQIK	1.00	41.2	−0.5	1143.6	17891	
DLRCVNMVAD	LWHAPAPK	0.82	32.4	−1.5	1034.5	17877	
DLRCVNMVAD	LWHAPAPK	0.91	30	1.9	1034.5	17877	
TNTIGSSEAC	MLGGMAMK	1.00	42	−0.8	953.4	17877	
TNTIGSSEAC	MLGGMAMK	1.00	38.7	0.5	953.4	17877	
AELDDIFSVQ	NLMHPAYK	0.98	32.4	−1.6	1088.5	87082	
AELDDIFSVQ	NLMHPAYK	0.93	38.8	−1.7	1088.5	87082	
ALNMIDYGLD	NLPGGPL	0.68	26.4	−2.2	754.4	17890	
GLEEIGVVCQ	QAGGHAAFVDAGK	1.00	78.5	−1.8	1343.6	87082	
GLEEIGVVCQ	QAGGHAAFVDAGK	1.00	77.7	−1.1	1343.6	87082	
DGLEEIGVVC	QQAGGHAAFVDAGK	1.00	86.9	−0.8	1471.7	87082	
DGLEEIGVVC	QQAGGHAAFVDAGK	1.00	86.9	−0.5	1471.7	87082	
GLKQCKANPW	QQFAETHNK	1.00	44.5	−0.7	1217.5	17871	
GLKQCKANPW	QQFAETHNK	0.98	39.3	−0.8	1217.5	17871	
EHDPMEIWAT	QSSTLVEVLAK	0.93	35.3	0.2	1289.7	17903	
EHDPMEIWAT	QSSTLVEVLAK	0.97	39.3	−2.2	1289.7	17903	
AESVKNIFGY	QYTIPTHQGR	1.00	46.3	−1.7	1287.6	87082	
AESVKNIFGY	QYTIPTHQGR	1.00	46.3	−1.0	1287.6	87082	
PYIVATITSN	SAGGQPVSLANLK	0.94	35.1	0.3	1356.7	87082	
PYIVATITSN	SAGGQPVSLANLK	1.00	42	−2.4	1356.7	87082	
ARLMATMKEA	SAGKLVR	0.98	32.7	0.2	845.5	17904	
YAGQDIVSNA	SCTTNCLAPLAK	1.00	45.4	−6.1	1450.7	17880	
RKIKAAQYVA	SHPGEVCPAK	0.99	43.7	−0.6	1196.5	17868	
RKIKAAQYVA	SHPGEVCPAK	1.00	38.5	−0.8	1196.5	17868	
RKIKAAQYVA	SHPGEVCPAK	1.00	47.8	−0.8	1196.5	17868	
PFACIAAGIA	SLWGPAHGGANEAALK	0.96	39.7	1.3	1693.8	17869	
PFACIAAGIA	SLWGPAHGGANEAALK	0.97	43.9	−1.7	1693.8	17869	
HDPMEIWATQ	SSTLVEVLAK	0.95	35.2	−0.8	1161.6	17903	
HDPMEIWATQ	SSTLVEVLAK	0.99	38.2	0.0	1161.6	17903	
RTMACGIAGL	SVAADSLSAIK	1.00	56.7	−0.1	1176.6	17871	
RTMACGIAGL	SVAADSLSAIK	1.00	65.1	4.3	1176.6	17871	
RTMACGIAGL	SVAADSLSAIK	1.00	50.7	−5.6	1176.6	17871	
CKKNPQKNLY	TFKNQASNDLPN	1.00	29.6	−0.6	1463.7	87082	
CKKNPQKNLY	TFKNQASNDLPN	0.98	26.7	1.4	1463.7	87082	
KEASAGKLVR	TLAAVRDAKEAA	0.99	38.5	−0.6	1330.7	17904	X
KEASAGKLVR	TLAAVRDAKEAA	0.99	43.5	−1.0	1330.7	17904	X
KEASAGKLVR	TLAAVRDAKEAA	0.99	35.7	−0.8	1330.7	17904	X
KEASAGKLVR	TLAAVRDAKEAA	0.97	34.6	0.1	1330.7	17904	X
EEKARGITIN	TSHVEYDTPTR	1.00	46.4	0.9	1392.6	17897	
EEKARGITIN	TSHVEYDTPTR	1.00	46.3	−1.4	1392.6	17897	
TTGEHEVSFQ	VHSEVFAK	0.83	27.9	−0.7	1031.5	17906	
QQAVAAHKFN	VLASQPADFDR	0.81	25.5	−0.7	1305.6	17901	
PDEAVAIGAA	VQGGVLTGDVK	0.99	40.8	−0.2	1187.6	17861	
PDEAVAIGAA	VQGGVLTGDVK	0.99	46.6	−1.8	1187.6	17861	
MVMPGDNIKM	VVTLIHPI	0.83	26.1	0.3	978.6	17897	
MVMPGDNIKM	VVTLIHPI	0.99	29.7	−1.4	978.6	17897	

Nonprime-side sequences are reported up to the position P10. X denotes ambiguous nonprime side residues. A small number of cleavage sites have P1 residues, which correspond to a potential cleavage-site of the digestion protease used for library generation. These cases might be related to incomplete amine protection during library preparation. Hence, prime-side cleavage products with such amino-termini are omitted from the PICS analysis, as well as prime-side cleavage products with amino termini that correspond to a protein amino terminus. One exemplary protein is stated for each identified peptide sequence. X! Tandem was applied to assign peptides to tandem mass spectra and the X!Tandem hyperscore is reported together with the PeptideProphet probability score. C signifies carboxyamidomethylated cysteine and K signifies di-methylated lysine. Some peptides were identified repeatedly by LC-MS/MS; however, the set of sequences was rendered nonredundant before analysis in the form of sequence logos.