Transcriptional knockdown in pneumococci using CRISPR interference

Morten Kjos

Faculty of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, P. O. Box 5003, 1432 Ås, Norway

Running Head: CRISPRi in pneumococcus

Abstract

Sequence specific knockdown of gene expression using CRISPR interference (CRISPRi) was recently developed for *Streptococcus pneumoniae*. By co-expression of a catalytically inactive Cas9-protein (dCas9) and a single guide RNA (sgRNA), CRISPRi can be used to knock down transcription of any gene of interest. Gene specificity is mediated by a 20 bp sequence on the sgRNA, and new genes can be targeted by replacing this 20 bp sequence. Here, a protocol is provided for design of sgRNAs and construction of CRIPSRi strains in *S. pneumoniae*, based on the vectors published by Liu et al. [1].

Key Words

CRISPRi, dCas9, knockdown, sgRNA, inverse PCR

1. Introduction

Functional studies of essential genes in bacteria rely on construction of depletion or knockdown mutants, since inactivation of such genes are lethal. Knockdown or depletion strains can be constructed by conventional approaches, where an inducible copy of the gene is introduced in an ectopic locus in the chromosome (or on a plasmid), followed by deletion of the native gene. Expression of the essential gene can then be titrated by different inducer concentrations. Another possibility to knock down expression of genes and operons is to utilize the CRISPR/Cas9-based technology known as CRISPR interference (CRISPRi) [2]. With CRISPRi, a catalytically dead Cas9 protein (dCas9) is used to selectively knock down the expression of a gene of interest. Unlike the wild-type Cas9 nuclease, dCas9 does not cleave DNA, but the DNA-binding capability is still intact. A single guide RNA (sgRNA) [3], containing a gene-specific base-pairing region and a structured region for interaction with dCas9, is designed to target the gene of interest. Upon co-expression, the dCas9-sgRNA complex binds DNA close to the 5' end of the gene and serves as transcriptional roadblock for RNA polymerase, thereby downregulating transcription (**Fig. 1**). The level of knockdown with CRISPRi can also be titrated by expressing one of the components (dCas9 or sgRNA) from an inducible promoter. A major advantage of CRISPRi over conventional construction of knockdown strains, is that new genes can be targeted by a single cloning step; only the 20 nt base pairing region of the sgRNA constructs needs to be changed to knock down a gene of interest. This allows for construction of large libraries of sgRNA strains [1, 4, 5]. On the other hand, a disadvantage with the system is the polar effects when targeting genes within an operon. Since the mechanism involves blocking transcription, all genes downstream of a target gene in an operon will likely be affected [1].

Streptococcus pneumoniae does not contain an endogenous CRISPR/Cas system [6, 7], and CRISPR/Cas9 can therefore be harnessed for different purposes in this bacterium. Liu et al. [1] recently developed an inducible CRISPRi for *S. pneumoniae*; vectors for chromosomal integration of sgRNA and *dcas9* constructs by double crossover were made. The CRISPRi system was shown to be specific and titratable [1]. In the same work, a library of CRISPRi strains targeting all essential genes in *S. pneumoniae* strain D39 was constructed. The collection comprises approximately 350 strains, which were all phenotypically characterized and eventually used for identifying the function of uncharacterized genes involved in cell wall synthesis and competence development [1]. Also worth noting, CRISPR/Cas9 has also been harnessed for other purposes in *S. pneumoniae*, including introduction of double strand breaks in DNA in strain D39 [8], and to mutagenize genes in *S. pneumoniae* R6 [7].

In this chapter, it will be explained how an sgRNA should be designed to effectively and specifically target a gene of interest and how the vectors designed by Liu et al. [1] can be used to construct strains for CRISPRi knockdown.

2. Materials

<u>Related to design and construction of new sgRNA plasmids.</u>

- Genome sequence of your pneumococcal strain
- sgRNA template plasmid and sequence: pPEPX-P3-sgRNAluc (Addgene #85590)
- Universal, 5' phosphorylated reverse primer (5'-TATAGTTATTATACCAGGGGGACAGTGC-3') (*see* **Note 1**).
- Designed reverse sgRNA primer containing the 20 bp base pairing sequence
- Reagents for high fidelity PCR (Phusion polymerase, dNTPs)
- Equipment for agarose gel electrophoresis
- PCR purification kit

- DpnI enzyme and buffer
- T4 Ligase
- *Escherichia coli* cloning host (e.g. DH5α)
- LB broth (10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract)
- LA agar plates (10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract, 1.5 % agar) with 100 µg/ml spectinomycin
- Plasmid purification kit
- Sequencing primer (Ppepx_Seq_R; 5'-CGAGGGATTTGGTGATTCTTCTT-3')

Related to the construction of CRISPRi strains

- Competent pneumococcal strain and transformation protocol
- Plasmid for integration of constitutively expressed *lacI*: pPEPY-PF6-lacI (Addgene #85589)
- Plasmid for integration of IPTG-inducible *dcas9*: pJWV102-PL-dCas9 (Addgene #85588).
- C+Y medium [9, 10]
    - C+Y medium contains (total 110 mL): 100 mL PreC, 2.5 mL Adams III, 2.5 mL 10 % yeast extract, 1 mL 8 % BSA, 1.5 mL 2 % sodium pyruvate, 1 mL 20 % glucose, 0.5 mL 2 mg/mL uridine, 0.5 mL 2 mg/mL adenosine, 0.1 mL 0.4 mM $MnCl_2$, 0.073 mL 3 % glutamine, 0.327 mL 0.3 M sucrose. pH can be adjusted with HCl.
    - PreC contains 8.5 g/L K_2HPO_4, 5 g/L casein hydrolysate, 2 g/L sodium acetate, 11.25 mg/L cysteine, 6 mg/mL tryptophane.
    - Adams III contains 24 mg/L biotin, 24 mg/L nicotinic acid, 28 mg/L pyridoxine HCl, 96 mg/L calcium pantothenate, 26 mg/L thiamine HCl, 11 mg/L riboflavin, 20 mg/L $FeSO_4 \cdot 7H_2O$, 20 mg/L $CuSO_4 \cdot 5H_2O$, 20 mg/L $ZnSO_4 \cdot 7H_2O$, 8 mg/L $MnCl_2 \cdot 4H_2O$, 20 g/L $MgCl_2 \cdot 6H_2O$, 1,75 g/L L-asparagine, 200 mg/L choline, 0.5 g/L $CaCl_2$.

- Todd Hewitt agar plates (Todd Hewitt broth supplemented with 1.5 % agar) with 40 µg/ml gentamycin, 1 µg/ml tetracycline or 100 µg/ml spectinomycin (see **Note 2**)
- Isopropyl β-D-1-thiogalactopyranoside (IPTG)

3. Methods

3.1 Design and construction of new sgRNAs for CRISPRi

The sgRNA construct consists of a 20 nt base-pairing region, a Cas9 handle and transcriptional terminator (**Fig. 1**). The latter two remain constant for all sgRNA constructs, while the base pairing region ensures the gene specificity. When designing a novel sgRNA to target a gene of interest, there are several important criteria which needs to be taken into consideration:

1. The 20 nt base pairing region should preferable bind to the non-template DNA strand close to the 5' end of the gene or the promoter region to obtain optimal efficiency. Binding to the template DNA strand has been shown to be less efficient [2].
2. The base pairing region needs to be located adjacent to the protospacer adjacent motif (PAM) sequence of Cas9 [3]. The constructs designed by Liu et al. [1] utilize Cas9 from *Streptococcus pyogenes*, which has a 5'-NGG-3' (any nucleotide followed by two guanosine nucleotides) PAM-sequence [6]. Partial targeting has also been observed for PAM-site 5'-NAG-3' [7]. Only sequences next to 5'-NGG-3' should be selected as base pairing regions.
3. To ensure specific knockdown of the gene of interest, it is critical to ensure that the sgRNA does not target other genes in the genome. The 12 nts proximal to the PAM sequence (and thus the Cas9-handle in the sgRNA sequence) are most important for specificity. This sequence is known as the "seed sequence" and 1-2 differences here will dramatically reduce the binding efficiency of the sgRNA [2, 11]. BLAST searches should be performed to ensure that there are no off-target binding sites.

4. The secondary structure of the sgRNA needs to be intact for dCas9 to bind to the sgRNA-DNA complex. The folding of a newly designed sgRNA should therefore be checked using a secondary structure folding algorithm such as RNAfold from the ViennaRNA package [12] (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) and compared to a functionally verified sgRNA.

As an example, it will here be explained how to design and construct an sgRNA plasmid targeting pneumococcal *pbp2a* gene (SPV_1821, SPD_1821, spr1823), encoding a bi-functional penicillin binding protein [13, 14], using inverse PCR (*see* **Note 3**).

1. Search the transcribed sequence (non-template strand) of the *pbp2a* gene (**Fig. 1**) for PAM-sequences 5'-CCN-3' (reverse complement to 5'-NGG-3' on the template strand). The 20 nt downstream of this PAM is a potential base pairing sequence (**Fig. 1**) (*see* **Note 3**).
2. Perform a BLAST search against the full genome using the PAM-proximal 12 bp (seed sequence) as query. Discard any sgRNA whose seed sequence maps to a secondary site in the genome next to a PAM sequence.
3. Design the gene-specific primer for inverse PCR by adding the reverse complement of the 20 nt base-pairing sequence from the non-template strand to the annealing part of the forward primer (5'-GTTTAAGAGCTATGCTGGAAACAGC-3'). The resulting, gene specific primer for *pbp2a* will thus be: 5'-TTTTCGAATCGGACCTACTTGTTTAAGAGCTATGCTGGAAACAGC-3' (base-pairing sequence is underlined) (**Fig. 1**).
4. To ensure that the secondary structure of the full sgRNA is not influenced by the base pairing sequence, compare the secondary structure prediction of the new full sgRNA sequence (base pariring region + Cas9 handle + transcriptional terminator; (i.e., sgRNA(pbp2a);

   ATTTTCGAATCGGACCTACTTGTTTAAGAGCTATGCTGGAAACAGCATAGCAA

GTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTG CTTTTTTTGAAGCTTGGGCCCGAACAAAAACTCAT) with that of a functionally verified sgRNA (i.e., sgRNA(luc); ATAGAGGATAGAATGGCGCCGTTTAAGAGCTATGCTGGAAACAGCATAGCAA GTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTG CTTTTTTTGAAGCTTGGGCCCGAACAAAAACTCAT) using RNAfold from the ViennaRNA package (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) [12].

5. Order the verified oligo for inverse PCR.
6. Amplify new sgRNA vector using the following reaction:

	Volume (µl)
Phusion polymerase	0.5
HF buffer (10x)	10
dNTPs (2.5 mM each)	4
Universal primer (100 µM)	0.25
Specific primer (100 µM)	0.25
Template plasmid (100 ng/µl)	0.5
dH2O	34.5
Total	*50*

The following cycling conditions are used:

Temperature (°C)	*Time*	
98	60 sec	Initial denaturation
98	20 sec	25 cycles
60	30 sec	
72	30 sec/Kbp	
72	10 min	Final elongation

7. Cast a 1 % agarose gel and perform electrophoresis. Purify the PCR product from gel using a PCR purification kit. Elute the purified product PCR product in 20-30 µl elution buffer from the purification kit.
8. Set up two identical DpnI digestion reactions, to degrade template DNA (*see* **Note 5**).

	Volume (µl)

Purified PCR product	1
DpnI	1
Buffer (10 x)	2
dH_2O	16
Total	*20*

Incubate the reaction at 37°C for 2 hours.

9. To circularize the plasmid, set up ligation reactions with the DpnI-digested PCR reactions without further purification. Include an extra reaction without ligase as a control.

	Inverse PCR reaction (volume in µl)	*Control reaction (volume in µl)*
Digestion mix	10	10
Ligase buffer (10x)	2	2
T4 DNA Ligase	1	0
dH2O	7	8
Total	*20*	*20*

Incubate at room temperature for 2 hours or at 16°C overnight.

10. Transform both ligation reactions to *E. coli* using conventional heat-shock procedures [15]. Plate out on LA plates containing 100 µg/ml spectinomycin for selection and incubate at 37°C overnight (*see* **Note 4**).
11. The number of colonies of the plate with circularized plasmid should far exceed the number of colonies on the control plates (*see* **Note 6**). Pick a colony in LB with 100 µg/ml ampicillin and grow up 12-18 hours.
12. Store the strain as freeze stock and isolate plasmid. Verify correct sgRNA by Sanger sequencing using sequencing primer (Ppepx_Seq_R; 5’-CGAGGGATTTGGTGATTCTTCTT-3’).

3.2 CRISPRi strain construction

The CRISPRi system developed by Liu et al. [1] relies on LacI-based IPTG-inducible expression of dCas9. In addition to the *dcas9* and sgRNA constructs, a *lacI* gene thus needs to be integrated into the chromosome of a strain to be used for CRISPRi. All constructs are

available on plasmids via Addgene, and they all integrate into non-essential chromosomal loci by double crossover (*see* **Note 7**).

1. To introduce a constitutively expressed *lacI* into the pneumococcal genome (Note 1), transform the plasmid pPEPY-PF6-*lacI* into the pneumococcal strain. The PF6-*lacI* construct will integrate into the *prs1* locus [1, 16, 17]. Select correct transformants by plating out on TH plates containing 40 µg/ml gentamycin.
2. Pick colonies in C+Y medium containing 40 µg/ml gentamycin (*see* **Note 8**) and grow until OD_{600} =0.4 before the culture is stored as freeze stocks. Verify correct integration by PCR.
3. Next, transform the strain with plasmid pJWV102-PL-dCas9 [1], which integrates the LacI-dependent P_{lac}-*dcas9* construct into the *bgaA*-locus. Select transformants on TH plates with 1 µg/ml tetracycline.
4. Pick colonies in C+Y medium containing 1 µg/ml tetracycline (*see* **Note 8**) and grow until OD_{600} =0.4 before the cultures are stored as freeze stocks. Verify correct integration by PCR. The resulting strain, carrying both constitutive *lacI* and IPTG-inducible *dcas9*, can be used to introduce different sgRNAs.
5. Finally, transform the strain with the constructed sgRNA plasmid (pPEPX-sgRNA). The constitutively expressed sgRNA will integrate into the region between the genes *amiF* and *treR*. Transformants are selected on TH agar plates with 100 µg/ml spectinomycin.
6. Pick colonies in C+Y medium containing 100 µg/ml spectinomycin (*see* **Note 8**) and grow until OD_{600} = 0.4 before the cultures are stored as freeze stocks. Verify correct integration by PCR.
7. The resulting strain is ready for CRISPRi knockdown experiment using an assay of choice. For example, grow the CRISPRi strain in C+Y medium without antibiotics until OD_{600} = 0.4. Then, dilute the culture in C+Y medium with 1 mM IPTG for maximum knockdown (*see* **Note 9**).

4. Notes

1. Phosphorylated primer can be ordered directly from oligo providers.
2. Columbia blood agar or other media are also possible to use. When plating pneumococci on top of the agar, the plates need to be incubated in anaerobic environment (anaerobic jars or 5 % CO_2 incubators).
3. It is here described how to use inverse PCR to create new sgRNA plasmids (introduce new 20 bp sequences in the vector). In addition to inverse PCR, Liu et al. [1] describes how to use infusion cloning for this purpose.
4. Automatic searches for base pairing sequences can also be done using available softwares, such as CRISPR Primer Designer [18].
5. One of the parallel reactions will be used as negative control in the ligation reaction.
6. A large number of colonies on the control plate suggests that the DpnI reaction, which should degrade the template plasmids, has not worked properly. The DpnI treatment then needs to be repeated.
7. All vectors of the CRISPRi system constructed by Liu et al. [1] contain sequences with homology to the chromosome of *S. pneumoniae* D39, which allows them to integrate into non-essential chromosomal loci by double crossover. For utilization in other pneumococcal strains, the sequences of the homology regions should be similar to D39, and this should be checked before starting the experiment.
8. Instead of picking and growing the colonies in liquid medium containing antibiotics, the colonies can also be re-plated on antibiotic plates and incubated over-night. Re-plated colonies can then be picked and grown in liquid medium without antibiotics.
9. The CRISPRi system is titratable, and the level of knockdown can be adjusted by reducing the IPTG concentrations [1].

5. References

1. Liu, X., Gallay, C., Kjos, M., Domenech, A., Slager, J., van Kessel, S.P., Knoops, K., Sorg, R.A., Zhang, J.R., and Veening, J.W. (2017). High-throughput CRISPRi phenotyping identifies new essential genes in *Streptococcus pneumoniae*. Mol Syst Biol 13, 931.
2. Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. Cell 152, 1173-1183.
3. Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337, 816-821.
4. Larson, M.H., Gilbert, L.A., Wang, X., Lim, W.A., Weissman, J.S., and Qi, L.S. (2013). CRISPR interference (CRISPRi) for sequence-specific control of gene expression. Nat Protoc 8, 2180-2196.
5. Peters, J.M., Colavin, A., Shi, H., Czarny, T.L., Larson, M.H., Wong, S., Hawkins, J.S., Lu, C.H., Koo, B.M., Marta, E., et al. (2016). A comprehensive, CRISPR-based functional analysis of essential genes in bacteria. Cell 165, 1493-1506.
6. Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucl Acids Res 41, 7429-7437.
7. Jiang, W., Bikard, D., Cox, D., Zhang, F., and Marraffini, L.A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. Nat Biotechnol 31, 233-239.
8. van Raaphorst, R., Kjos, M., and Veening, J.W. (2017). Chromosome segregation drives division site selection in *Streptococcus pneumoniae*. Proc Natl Acad Sci U S A 114, E5959-E5968.

9. Martin, B., Garcia, P., Castanie, M.P., and Claverys, J.P. (1995). The *recA* gene of *Streptococcus pneumoniae* is part of a competence-induced operon and controls lysogenic induction. Mol Microbiol 15, 367-379.
10. Lacks, S., and Hotchkiss, R.D. (1960). A study of the genetic material determining an enzyme in Pneumococcus. Biochim Biophys Act 39, 508-518.
11. Hawkins, J.S., Wong, S., Peters, J.M., Almeida, R., and Qi, L.S. (2015). Targeted Transcriptional Repression in Bacteria Using CRISPR Interference (CRISPRi). Methods Mol Biol 1311, 349-362.
12. Lorenz, R., Bernhart, S.H., Honer Zu Siederdissen, C., Tafer, H., Flamm, C., Stadler, P.F., and Hofacker, I.L. (2011). ViennaRNA Package 2.0. Algorithms Mol Biol 6, 26.
13. Paik, J., Kern, I., Lurz, R., and Hakenbeck, R. (1999). Mutational analysis of the *Streptococcus pneumoniae* bimodular class A penicillin-binding proteins. J Bacteriol 181, 3852-3856.
14. Hoskins, J., Matsushima, P., Mullen, D.L., Tang, J., Zhao, G., Meier, T.I., Nicas, T.I., and Jaskunas, S.R. (1999). Gene disruption studies of penicillin-binding proteins 1a, 1b, and 2a in *Streptococcus pneumoniae*. J Bacteriol 181, 6552-6555.
15. Sambrook, J.F., and Russell, D.W. (2001). Molecular Cloning: A Laboratory Manual, Volume 1,2 and 3, 3rd Edition, (Cold Spring Harbor Laboratory Press).
16. Sorg, R.A., Kuipers, O.P., and Veening, J.W. (2015). Gene expression platform for synthetic biology in the human pathogen *Streptococcus pneumoniae*. ACS Synth Biol 4, 228-239.
17. Sorg, R.A. (2016). Engineering approaches to investigate pneumococcal gene expression regulation and antibiotic resistance development. PhD thesis. (University of Groningen).
18. Yan, M., Zhou, S.R., and Xue, H.W. (2015). CRISPR Primer Designer: Design primers for knockout and chromosome imaging CRISPR-Cas system. J Integr Plant Biol 57, 613-617.

Figure legends

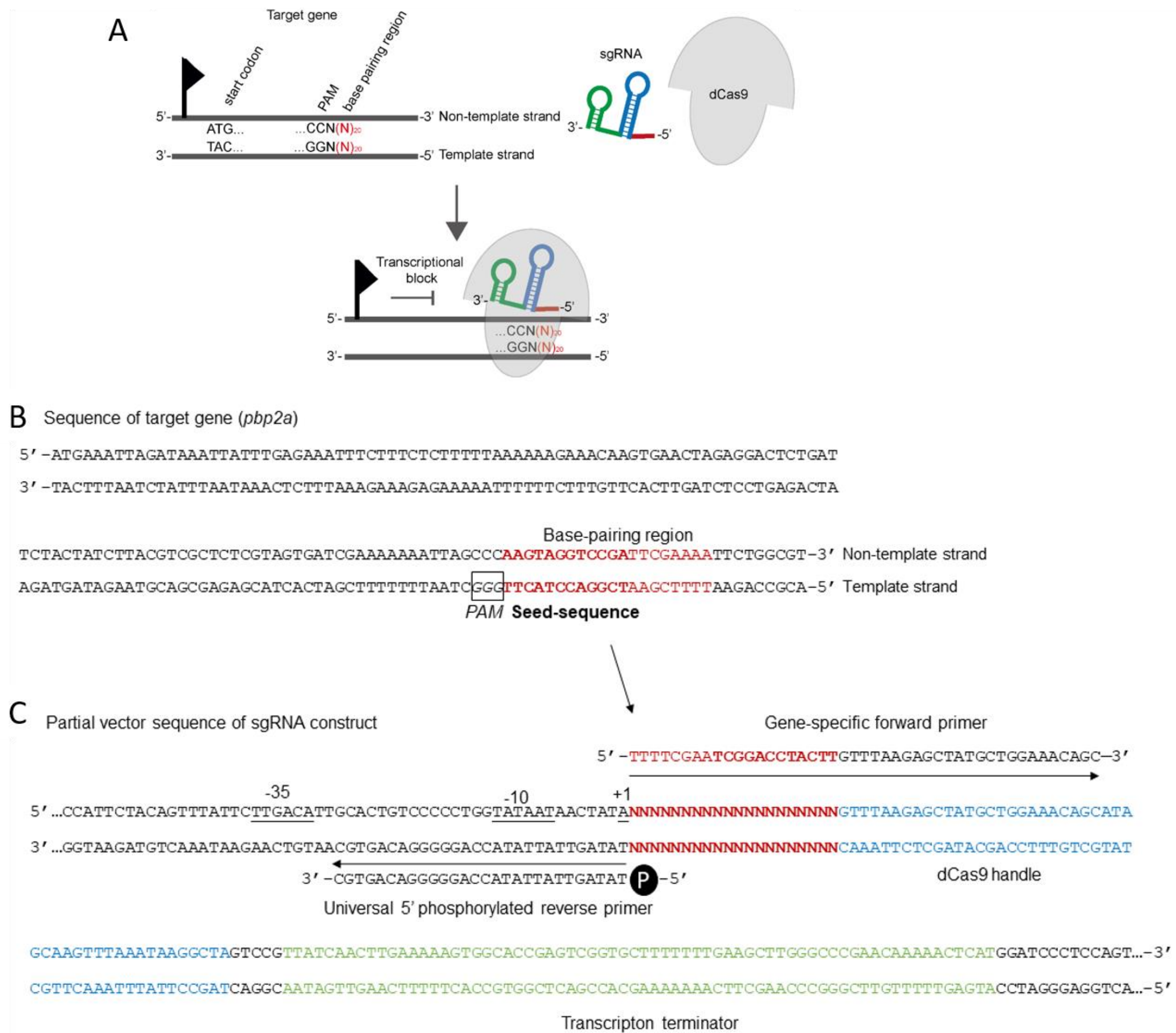

Figure 1. Overview of the CRISPRi system.

A. Schematic overview of CRISPRi. The start codon, PAM-sequence and base pairing region of the target genes are shown as well as the sgRNA and dCas9 protein. The promoter is shown as a flag. When sgRNA binds to the non-template strand, the DNA-sgRNA-dCas9 complex form a transcriptional roadblock causing knockdown of expression of the target gene. **B and C**. Inverse PCR primer design to construct novel sgRNAs, using the pneumococcal gene *pbp2b* as an example. **B**. The beginning of the *pbp2b*-encoding sequence including start codon ATG. The first PAM site is indicated (box), as well as the base-pairing region (red). The 12 bp seed sequence of the base pairing region is shown in bold. **C**. Construction of the sgRNA vector.

The base-pairing region is introduced as an overhang on the forward primer, which anneals in the dCas9 handle region. The reverse primer is universal and phosphorylated in the 5'end to allow ligation of the linearized vector.