

Article

An Insight into Strain-Specificity of *Streptomyces chrestomyceticus* ADP4 and Identification of a Novel Peptide with Potential Antiviral Activities Against Significant Human Viruses, Including SARS-CoV2, HCV, and HIV

Varsha Verma ¹, Medicherla Krishna Mohan ²  and Ashok K. Dubey ^{3,*} 

¹ Department of Applied Sciences (Biotechnology), NSIT, Delhi University, New Delhi 110007, India; varsha92igib@gmail.com

² School of Life and Basic Sciences, Jaipur National University, Jaipur 302017, India; krishna.mohan@jnujaipur.ac.in

³ Department Biological Sciences and Engineering, Netaji Subhas University of Technology, New Delhi 110078, India

* Correspondence: akdubey@nsut.ac.in or adubey.nsit@gmail.com

Abstract

This study aimed to unravel the genomic uniqueness of *Streptomyces chrestomyceticus* ADP4 using whole-genome sequence analysis, with a focus on identifying strain-specific genes/proteins associated with a novel therapeutic source. The genome of the strain ADP4 was sequenced and assembled to a total size of 9.64 MB. A total of 8378 coding regions were identified. Strain ADP4 was found to be clustered into a clade of the species *S. chrestomyceticus*. Fifty-one biosynthetic gene clusters were predicted in the genome of the strain ADP4, and three of them were common to all the strains of *S. chrestomyceticus*. A comparative metabolic profile of *S. chrestomyceticus* revealed a unique metabolic protein, supporting strain-level variations. Comparative genome analysis led to the identification of the genomic sequences that were specific to the strain ADP4. These strain-specific unique sequences of ADP4 were identified across the available data, underscoring their distinct genetic identity. Among these eight functionally uncharacterized hypothetical proteins (HPs), only two could be assigned with functional attributes, wherein one of them, HP2, was ascertained to be a peptide with possible antiviral activity, underscoring its potential as a novel drug candidate for antiviral therapy. The structural validation and peptide–protein molecular docking have evidently demonstrated antiviral activity of HP2 against significant human viral pathogens, for example, HIV, SARS-CoV2, HCV, ZIKV, JEV, and DENV.

Keywords: *Streptomyces chrestomyceticus*; strain specificity; genome mining; comparative genomics; antiviral peptide (AVP); molecular docking



Academic Editor: Michal Letek

Received: 11 October 2025

Revised: 18 November 2025

Accepted: 24 November 2025

Published: 26 November 2025

Citation: Verma, V.; Mohan, M.K.; Dubey, A.K. An Insight into Strain-Specificity of *Streptomyces chrestomyceticus* ADP4 and

Identification of a Novel Peptide with Potential Antiviral Activities Against Significant Human Viruses, Including SARS-CoV2, HCV, and HIV. *Microbiol. Res.* **2025**, *16*, 249. <https://doi.org/10.3390/microbiolres16120249>

Copyright: © 2025 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Bacteria belonging to the actinomycetes group, particularly the genus *Streptomyces*, have long been known as key sources of therapeutic compounds [1]. In view of their proven potential to produce bioactive metabolites with unique chemical diversity and wide-ranging functions, they have become a preferred choice for the discovery of new drugs [2]. In the present era of emerging and emerged drug-resistant pathogens, infectious diseases have posed a serious threat to human health globally, which needs to be addressed

with a sense of urgency [3]. Our endeavour in this regard involved bioprospecting of members of the genus *Streptomyces* in search of new and effective drugs against difficult-to-treat ESKAPE pathogens [4]. During the pursuit of this work, we happened to isolate an actinobacterium, identified as *S. chrestomyceticus* strain ADP4, which proved to be a prolific producer of novel and potent antimicrobial compounds [5–12]. In view of its robust biosynthetic attributes, the strain ADP4 holds a good promise for the discovery and development of new drugs. This possibility assumes significance considering the growing menace of antimicrobial drug resistance (AMR) among pathogens. However, an understanding of the molecular machinery of the strain ADP4 would prove handy for the drug discovery process. Accordingly, molecular characterization of the strain ADP4 was undertaken from the genomic perspective. The genome sequence of the strain ADP4 was explored, along with related strains, to gain an insight into the distinctive features of its genome that might be associated with strain specificity.

Comparative pan-genome analysis provides an insight into the identification of strain-specific genes for genetic diversity, evolution and adaptation. Recent studies, for example, in the case of *Bacillus subtilis* [13], *Amycolatopsis* sp. [14], and *Streptomyces* spp. [15], have provided valuable insights into strain specificity up to the species/genus level. However, for a global view of strain specificity, the pan-genome needs to be explored beyond the genus level for the existing non-redundant database, since specificity, so determined, may unravel unique strain-specific biological functions, which was considered in the present study. In view of the antimicrobial potential of ADP4, exploring the genome could uncover the unique genetic signatures and functional traits that define strain-specificity. The genome was mined for the functional genes and BGCs. Comparative genomics of *S. chrestomyceticus* strains was performed to understand the genes and proteins that confer specificity to the strain ADP4. The present study, encompassing analyses of the uniqueness associated with the ADP4 genome and its functional characterization, could be expected to accelerate antimicrobial drug discovery.

2. Materials and Methods

2.1. Microorganism, Culture Conditions, and Genomic DNA Preparation

The actinomycete *S. chrestomyceticus* strain ADP4 used in the present work was reported earlier [5]. For genome isolation, a single colony from a 48 h old culture plate (nutrient agar, SRL, Mumbai, India) was inoculated in 20 mL of nutrient broth, which was incubated at 28 °C for 24 h under shaking at 150 rpm in an incubator shaker (ISF-1-W Kuhner, Birsfelden, Switzerland). The cell mass was recovered by centrifugation at 8000 rpm for 15 min. The genomic DNA (gDNA) from the cell mass was extracted by using the CTAB method [16]. Quality of the gDNA was checked by agarose gel electrophoresis (0.8% agarose) and by nano-drop measurements (Cary 60 UV-Vis, Agilent Technologies, Santa Clara, CA, USA).

2.2. Genome Sequencing, Assembly, and Annotation

Sequencing of the ADP4 gDNA was performed by the HiSeq System Illumina platform with 150 bp × 2 configurations. A standard pipeline of raw read analysis was applied. Quality of raw reads was evaluated using fastQC version 0.11.9 [17]. The raw reads, having low-quality bases from the ends of reads, were trimmed and the reads below the quality threshold were removed by using Trimmomatic (version 0.36) [18]. De novo assembly was performed by SPAdes (version 3.12.0) [19]. Genome completeness and contamination were evaluated with CheckM (v. 1.2.3) [20], and integrity was confirmed by using BUSCO v. 5.1.10 [21]. Structural and functional annotation of the genome was performed with DFAST v. 1.2.0 [22], PGAP v. 6.6 [23], the BPGA pipeline [24], COGs [25], and KEGG [26]. Predictive

identification of antibiotic resistance genes (ARGs) was performed by employing the CGE ResFinder 4.0 and CARD RGI (v. 4.0.0) software databases [27,28]. The default settings were used for all the tools that were deployed in the present study. The ADP4 genome was mined for BGCs by employing the AntiSMASH (version 7.0) bacterial database [29] using relaxed detection strictness and default parameters, enabling all extra annotation features. The biological function of all the predicted clusters and compounds was determined with the help of the PubChem database (<https://pubchem.ncbi.nlm.nih.gov/>; accessed on 7 June 2024) and reference literature. The conserved domains of the core genes were extracted using CDD [30].

2.3. Phylogenetic Analysis

Average Nucleotide Identity (ANI) values were calculated by EzBioCloud ANI Calculator [31]. Digital DNA–DNA hybridization (dDDH) value was determined by employing the Genome-to-Genome Distance Calculator (GGDC) [32]. The phylogenomic comparison of the ADP4 strain with related species/strains (Table S1) was performed by Type Strain Genome Server (TYGS) [32]. The constructed tree was visualized using ITOL (v. 7.1, <https://itol.embl.de/>) [33].

2.4. Analysis for the Specificity of ADP4

The comparative analysis based on orthologous gene clusters was performed with OrthoVenn3 [34]. Singletons specific to the strain ADP4 were selected. pBLAST (v. +2.15.0) was employed to select strain-specific sequences with <40% similarity and <80% QC. All the reference genome sequences of *S. chrestomyceticus*, as provided in Table S1, were extracted from the NCBI genome database (<https://www.ncbi.nlm.nih.gov/>). Structural and functional analyses of the strain-specific proteins were performed using various functional annotation tools. These included Pfam, InterPro, CATH, SUPERFAMILY, SMART, SCANPROSITE, CDD-BLAST, RSAT, and others. All these analyses were performed using the default parameters. The physicochemical parameters of selected proteins were theoretically determined by using Expasy's ProtParam server. Three-dimensional protein structure was predicted using the AlphaFold Protein Structure Database based on the amino acid sequence. A Ramachandran plot was generated using the ProCheck (<https://www.ebi.ac.uk/thornton-srv/software/PROCHECK/>, accessed on 15 May 2025) web server.

2.5. Antimicrobial Peptide Analysis

All strain-specific sequences were subjected to assessment of their antimicrobial potential by employing a multi-model approach, including the Collection of Anti-Microbial Peptides RF classifier (CAMPR3), SVM, and ANN, along with AMPLify. The sequences predicted as AMPs by two or more independent models were considered as potential AMP candidates with high confidence. The Database of Antimicrobial Activity and Structure of Peptides (DBAASP) was used to find antimicrobial targets. The amino acid sequences of DENZ2, HIV, HCV, JEV, and SARS-CoV2 were retrieved from the UniProt Knowledgebase (UniProtKB) and Protein Data Bank. The FASTA-formatted sequence was downloaded and used as the primary input for all downstream in silico analyses, including modelling and molecular docking. Protein–peptide docking was performed by using the HDock web server (<http://hdock.phys.hust.edu.cn/>). For each system, the receptor (protein target) and ligand (peptide) were submitted in PDB format after energy minimization and structural refinement. The docking process began with a template search from a curated interaction database. For those where no homologous complex was identified, global docking was executed to scan the full surface of the receptor to identify potential binding hotspots. For evaluation, each docking model was evaluated using three primary metrics: docking score

(kcal/mol), confidence score (0–1), and ligand RMSD (Å). The Computed Atlas Surface Topography of Proteins (CASTp) server was used for active-site prediction of proteins. Hydrogen bonding patterns and interface properties such as interface area, number of residues, salt bridges, disulfide bonds, and non-bonded contacts were also analyzed to assess interaction specificity and strength. Further, the pharmacokinetics (PK) and AD-MET properties of the peptides were assessed by using the open-source web-based tool SwissADME (www.swissadme.ch).

3. Results

3.1. Genome of *S. chrestomyceticus* Strain ADP4

The de novo-assembled genome of the strain ADP4, with a completeness of 99.24%, has been submitted to the NCBI Genome database under the accession number GCA_036562025.1. The genome size is approximately 9.6 Mb, with a GC content of 72% (Table 1). The NCBI database contained 12 genomes of the strains belonging to *S. chrestomyceticus*. The range of genome size varied from 9.31 Mb to 9.86 Mb, with GC contents of 72.05 + 0.05%. According to the NCBI Quality analysis (CheckM), the assemblies showed 95–99% completeness and 1–1.4% contamination. However, the strain JCM 4735, came up with a suppressed assembly and therefore could not be included for the comparative genome analysis. Based on the comparison of the genome annotation statistics (Table 1), strain ADP4 is catalogued in a group of *Streptomyces* with a larger genome size. There are 13 CRISPR loci in the ADP4 genome, which include Type I-E and Type V-F2 CRISPR-Cas systems.

Table 1. Comparative analysis of general feature annotation summary of *S. chrestomyceticus* strain ADP4 with closely related *S. chrestomyceticus*, performed using DFAST Prokaryotic genome annotation pipeline.

Strain Attribute	ADP4	NPDC020535	NPDC045187	NPDC045234	NPDC045757	NBRC 13444	TBRC 1925	NPDC003050	NPDC020200	GCAL-9	NPDC045451
Genome (Mb)	9.64	9.31	9.51	9.67	9.68	9.38	9.34	9.86	9.74	9.52	9.76
Number of CDSs	8378	8145	8299	8513	8465	8346	8096	8623	8430	8182	8589
GC content (%)	72.0	72.1	72.0	72.1	72.1	72.0	72.1	72.0	72.1	72.1	72.1
DNA scaffolds	68	35	447	83	72	1	66	52	67	246	66
Gap ratio (%)	0.002	0.003	0.005	0.008	0.004	0.0	0.0	0.0	0.005	0.0	0.004
Coding ratio (%)	86.2	87.6	83.3	86.4	86.4	87.2	87.3	86.1	87.0	84.8	86.0
N50 (bp)	345,478	693,272	77,237	254,561	254,010	377,418	597,445	668,145	329,609	84,471	393,997
Number of CRISPR loci	13	6	3	2	3	6	4	13	10	2	3
Number of rRNAs	8	2	2	2	3	21	2	2	2	4	2
Number of tRNAs	90	92	85	88	94	92	94	96	111	87	89

Gene ontology annotations covered biological processes (65%), molecular functions (47%), and cellular categories (59%) (Figure S1). Further, Clusters of Orthologous Groups (COGs) provided a rational classification of proteins encoded by the genome sequence. Each COG included proteins inferred to be orthologs (direct evolutionary counterparts). Out of all proteins encoded, 61% could be categorized into functional groups; the remaining

39% of genes were assigned as non-functional (Table S2). These genes were related to the following classes: carbohydrate transport and metabolism (4.45%), energy production and conversion (4.99%), and others. About 26% of the functionally categorized genes could not be assigned any function. A minority class, including only 1–2% of the genes, was found to be associated with cell motility, nucleotide transport and metabolism, lipid transport and metabolism, and secondary metabolite synthesis.

Analysis of the metabolic pathways led to the prediction of 2775 genes, which constituted 34.7% of CDSs. These genes were associated with 22 functional groups, including carbohydrate metabolism (12%), protein families of signalling and cellular processes (11.33%), amino acid metabolism (7.8%), metabolism of terpenoids and polyketides (2.27%), biosynthesis of other secondary metabolites (1.1%), xenobiotics biodegradation and metabolism (0.9%), and environmental information processing (8.5%), as shown in Figure S2. Out of all the identified genes, 4.68% remained unclassified. Comparative metabolic pathway analysis of *S. chrestomyceticus* strains revealed 14 genes that could be annotated and belonged to the metabolic functional category. The genes included in the categories were spoIVCA (site-specific DNA recombinase), dam (DNA adenine methylase), pks5 (polyketide synthase 5), vapB6S (antitoxin VapB6/11/15/18/21), vapC2S (ribonuclease VapC2/6/11/15/17/18/21/51), NHA1, SOD2 (sodium/hydrogen antiporter), PDFdef (peptide deformylase), thnH (4'-phosphopantetheine phosphatase), pnbA (para-nitrobenzyl esterase), ABC.CD.A and ABC.CD.P (putative ABC transport system permease protein), mannosyltransferase, lolD (lipoprotein-releasing system ATP-binding protein), and an uncharacterized protein. Among the above, pks5 and the uncharacterized protein were noted only in the strain ADP4. Blast search resulted in ADP4 as the first hit (100% identity), followed by *S. xanthochromogenes* (56% identity) and others with lower identity. Hence, the Pks5 gene appeared to be present only in the genome of the strain ADP4 among the genomes of the *S. chrestomyceticus* strains (Figure 1). In *Streptomyces* spp. and other antibiotic-producing bacteria, varying numbers of ARGs were identified [31]. ARGs belonging to the families, such as AAC(3)-VIIa, folA, Dfr, VanX/O/H, Tet, OtrB/C, S12p/10p, PgsA, OxyR, MurA, MtrA/B, LpqB, MprF, GdpD, and CmlV were predicted for the strain ADP4. Resistance genes predicted with more than 90% identity included aminoglycoside resistance (96.31%), tetracycline resistance (93.88%), and glycopeptide resistance (99.43%).

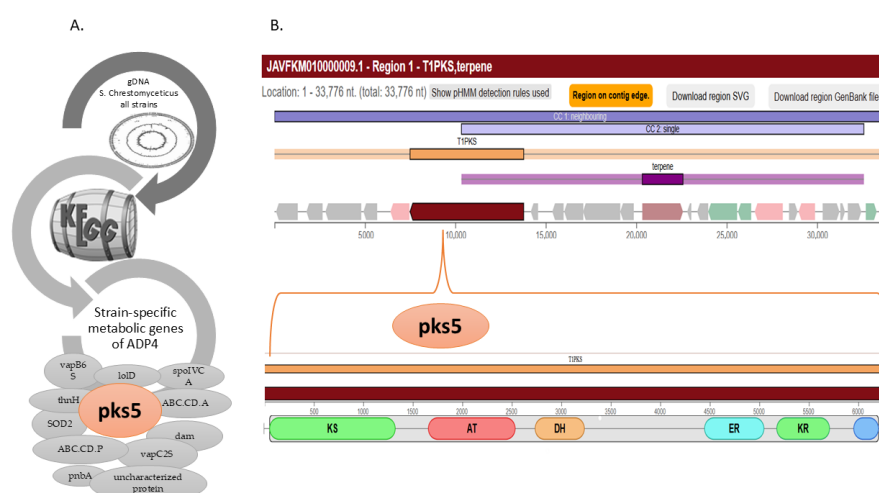


Figure 1. Strain-specific metabolic attributes of ADP4. (A) Fourteen ADP4-specific metabolic genes were identified by comparative KEGG analysis of the *S. chrestomyceticus* genomes. (B) The unique gene Pks5 corresponded with a biosynthetic gene cluster associated with the chromene and coumarin classes of metabolites.

3.2. Biosynthetic Gene Clusters of the Strain ADP4

Genome mining of the ADP4 strain with AntiSMASH 7.0 led to the prediction of 51 BGCs (Table S3), which included both overlapping and non-overlapping clusters with diversity in classes. As shown in Figure S3A, each cluster contained cluster-situated transcription-factor-binding sites (TFBSs). The TFBSs in NRPS and NRPS-like clusters have common sites that include DmdR1, BldD, DasR, and AbrC3, with significant confidence. Similarly, DmdR, NrdR, DasR, and BldR were observed for PKS and siderophore. The presence of a metal-responsive repressor as a TFBS was evident for terpene. Seven of the predicted BGCs of ADP4, which included cyclothiazomycin, EDHA, geosmin, ectoine, legonoxamine A/desferrioxamine B/legonoxamine B, antipain, and purincyclamide, showed 100% similarity with known BGCs. These BGCs are highly conserved in the genus *Streptomyces* and are present in almost all of their genomes. A total of 39% of BGCs showed > 50% similarity, 60% had lower similarity, and 1% remained unidentified. This suggested a high probability of novelty for the metabolites from such clusters. Comparative BGC analysis led to the identification of 19 common clusters among the strains of *S. chrestomyceticus* (Figure S3B). The closeness of the strain ADP4 and the strain NPDC045757 was evident from their BGC profiles, which shared 11 common clusters associated with known products (Figure S3B). Three mixed-orientation complete clusters, namely, antipain, legonoxamine/desferrioxamine, and purincyclamide were noted in all the strains of *S. chrestomyceticus*. Three among the predicted clusters were present only in the strain ADP4 and not in other strains of *S. chrestomyceticus*.

3.3. Genome-Based Phylogeny

Comparison of the phylogenomic metrics data and the ANI and dDDH values of the strain ADP4 with other strains of *S. chrestomyceticus* was undertaken (Table S4). Considering the threshold scores of 95% ANI and 70% dDDH values for species demarcation, the strain ADP4 belonged to *S. chrestomyceticus*. The reference genomes, used for phylogenomic analysis, were retrieved from the NCBI database. Overall features of the genomes considered for the analysis are listed in Table S4. Based on the 16S rDNA sequence phylogeny, a total of 51 strains of *Streptomyces* spp. were chosen for comparative analyses; 11 of these were strains of *S. chrestomyceticus*. A genome phylogeny was used to establish the phylogenetic relationships of *S. chrestomyceticus* strains. According to the phylogenetic tree, all the 11 strains of *S. chrestomyceticus* have a common ancestor root. Here, an internal node with bootstrap 100 has two clades. Each consists of two sister groups with three strains. Among these, the strains ADP4 and NPDC045757 (GCA 040646575.1) were found to be in the same branch, as presented in Figure 2, suggesting that both strains share a recent ancestor and are closely related sister strains. The strain ADP4 appeared as a new strain of *S. chrestomyceticus* among the currently known strains with whole-genome sequences available.

Tree scale: 0.01

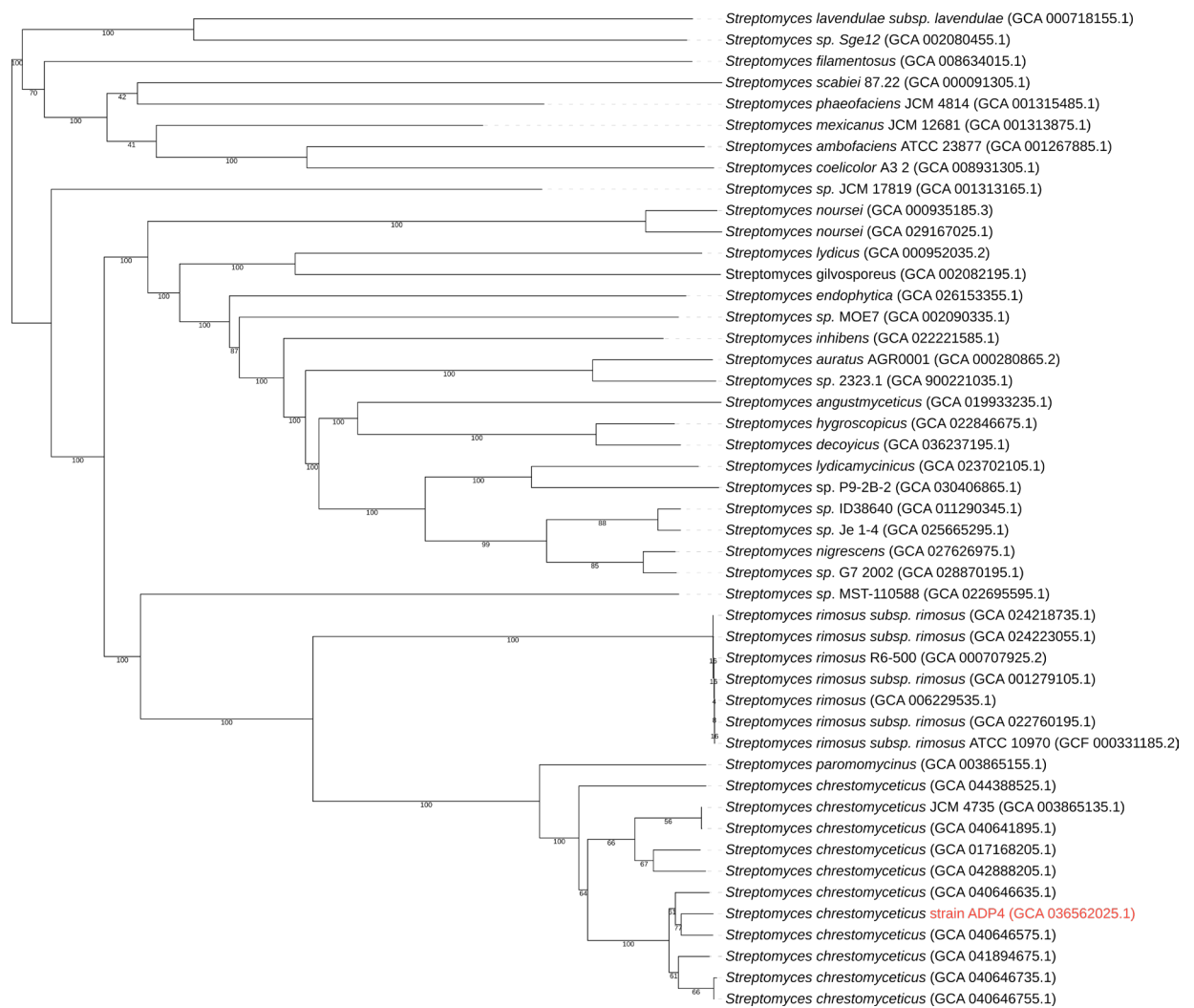


Figure 2. Phylogenetic tree of strain ADP4. Genome-sequence-based phylogenetic tree of strain ADP4 with closely related strains. The branches represent genome BLAST distance phylogeny (GBDP). The scale bar corresponds to 0.01 substitutions per nucleotide position. The position of the strain ADP4 in the phylogenetic tree is shown in red font.

3.4. Comparative Genomics

Pan-genome analysis of the strains of *S. chrestomyceticus* and *Streptomyces* spp. was performed to evaluate the genetic heterogeneity and phylogenetic relationships. The pan-genome of 11 strains of *S. chrestomyceticus* comprised an average of 13,500 gene sequences. This revealed the presence of a conserved core gene and an accessory and dispensable set of gene clusters (GCs). On average, core genes were found in > 38% of the genomes, and ~31% of the genes were dispensable genes. Moreover, >28% were designated as cloud genes, which included 2.5% GSs for the ADP4. About 56% of the cloud genes were designated hypothetical for the strain ADP4. These numbers were comparable with those for the closely related strains. The analysis categorized the protein sequence of *S. chrestomyceticus* strains into 7278 orthologous clusters with 154 singletons for strain ADP4 with major hypothetical proteins. These singletons could be involved in lineage-specific gene expansion. Among all orthologous clusters, 5985 were core GSs for *S. chrestomyceticus*, along with 2.2% singletons being strain-specific GSs. From both the analyses, it could

be observed that strain NPDC045757 possessed the highest number of clusters that were shared with the strain ADP4 among all the strains of *S. chrestomyceticus*.

All the ADP4-specific proteins were filtered using various platforms, as mentioned above. The strain-specific proteins with less than 40% identity and 80% query coverage were selected. Based on homology, these were found to be uncharacterized hypothetical proteins (HPs), which were functionally annotated by employing several effective bioinformatics tools. The results were evaluated in order to assign functions to HPs demonstrated by more than one tool, as listed in Table 2. The results revealed unsuccessful functional categorization for the proteins from HP3 to HP8 (Table 2), and thus, these were excluded from further studies.

Table 2. Functional and structural analysis of the hypothetical proteins using various bioinformatics tools and databases.

Sequence Name	NCBI Accession No.	Size (aa)	InterPro	DeepProLoc	TargetP	PSIPRED	CDD-BLAST	DeepTMHMM	Blastp Significant Hit (Cutoff 1 $\times 10^{-5}$)	DB with No Prediction for Any HP
HP1	WP_331790072.1	143	CC, SP, TM	Extracellular	SP	SP, extracellular	-	SP	-	
HP2	WP_331790253.1	132	IDR	Cytoplasmic	Other	Cytoplasmic	Partial match, 36%	-	HP hits with only 47% and less identity	
HP3	WP_331789357.1	60	IDR	Cytoplasmic Membrane	Other	Cytoplasmic	-	-	-	
HP4	WP_331786462.1	58	CC, SP, TM	Cytoplasmic	mTP	Cytoplasmic	-	-	Transposase	Pfam, CATH
HP5	WP_331786362.1	70	No match	Cytoplasmic Membrane	Other	Cytoplasmic	-	-	-	
HP6	WP_331788561.1	84	No match	Extracellular	Other	Extracellular	-	-	-	
HP7	WP_331789350.1	90	No match	Cytoplasmic	Other	Cytoplasmic	-	-	-	
HP8	WP_331786751.1	65	No match	Extracellular	Other	Extracellular	-	-	-	

“-” showed that no functional annotation could be achieved for the corresponding sequences.

3.5. Characterization of ADP4-Specific Sequences and the Novel Antiviral Peptide

A comparative analysis of CDSs from the genomes of different strains of *S. chrestomyceticus* led to the identification of eight unique CDSs which were present only in the strain ADP4 (Figure 3). Functional annotation of these CDSs could not yield any assignment to known functions attributed to proteins. Hence, these were designated as hypothetical proteins: HP1 to HP8. However, further studies were undertaken to explore if any application-related functions could be attributed to these HPs.

HP1 consisted of a 143 aa polypeptide whose physicochemical characteristics and cellular distribution were screened. The instability index (II), isoelectric point, GRAVY value, and molecular weight were computed to be 35.10, 9.40, -0.837 , 28.304, and 16.84 kDa, respectively (Table S5). Based on functional annotation, HP1 could be an extracellular protein with a signal peptide domain (Figure S4A). The homology modelling revealed poor alignment, with an RMSD value of 1.72, only 6.82% alignment, and a normalized score of -270.8 . The 3D structure of HP1 was created by Swiss-Model (Figure S4B), with a MolProbity score of 1.27 and zero rotamer outliers. The Ramachandran plot analysis of the likely structure of HP1 revealed 93.48% favourable regions (red dots), with 0% outliers. This indicated that conformation of HP1 was reliable, as shown in Figure S4C.

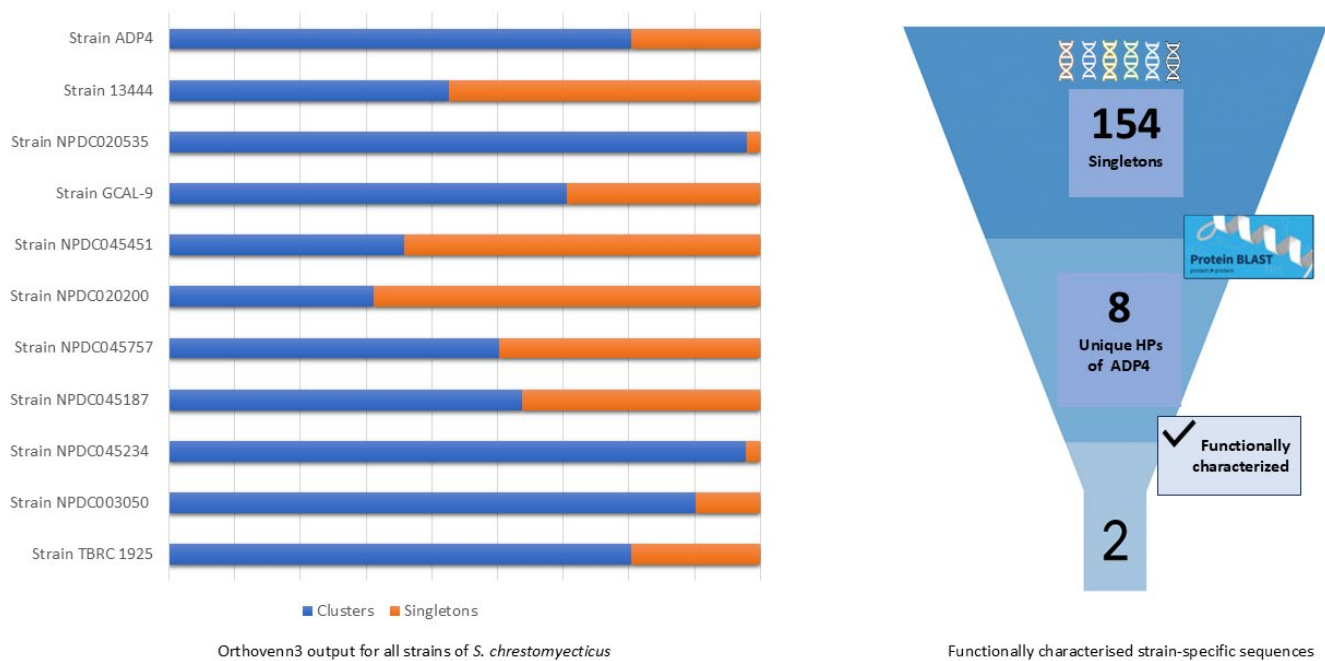


Figure 3. Comparative genome analysis. OrthoVenn3-based comparative genome analysis for identification of singletons, followed by BLASTp for strain-specific sequences.

HP2 was another unique protein of the strain ADP4 that comprised 132 aa. According to predictions, HP2 possessed a molecular weight of 14.58 kDa with a theoretical pI of 6.74. The GRAVY index was -0.603 , suggesting it to be hydrophilic and soluble (Table S5). Ala was found to be the most prevalent amino acid residue, followed by Arg and Gly, whereas there were no Pyl or Sec residues. It contained an equal number of positively (Asp + Glu) and negatively (Arg + Lys) charged amino acids, with 16 residues each. The molecular formula of HP2 was deduced as $C_{631}H_{999}N_{195}O_{192}S_6$, with 2023 atoms. Protein subcellular localization studies using CELLO, PSSpred, and DeepLocPro identified it as being a cytoplasmic protein with high confidence. To uncover the conserved domain and the potential function of the protein, various search tools were employed. InterPro suggested that it contained a consensus disorder prediction in positions 81–102. The NCBI CD search suggested that it contained a conserved curlin minor subunit CsgB (provisional) at positions 80–136 with a moderately significant e-score. As per MobiDB, CsgB often has partially distorted regions. These proteins are mostly found on the cell surface for adhesion and are extracellular in nature. This finding contradicts the physiochemical properties of HP2, i.e., soluble, hydrophilic, and cytoplasmic. As per ScanProsite, HP2 has no motifs other than motifs like phosphoserine and phosphothreonine, which are highly probable to occur. HP2 was not characterized using PANTHER, KEGG, or Pfam, and no interaction network could be generated. In the nr-search pBLAST of HP2, it showed similarity to *Streptomyces* hypothetical proteins, but with a percent identity of less than 47% at coverage of 70%. To observe the conserved and different residues among the homologous proteins, sequence alignment was performed, as presented in the Figure S5A. The low sequence identity to homologs supports the uniqueness of the protein, with divergent homologs in the genus. A phylogeny tree was constructed using the same data (Figure S5B). The secondary structure of HP2 included an alpha helix (41.67%), random coil (43.18%), and extended strand (15.15%), as in Figure 4A. Based on the Swiss-Model assessment, the 3D structure of HP2 showed a MolProbity of 2.33, a low clash-score, suggesting a good model structure (Figure 4B). The Ramachandran plot analysis of the likely structure of

HP2 showed 94.6% favourable regions (red dots), with 0% outliers. This indicated that conformation of HP2 was reliable, as shown in Figure 4C.

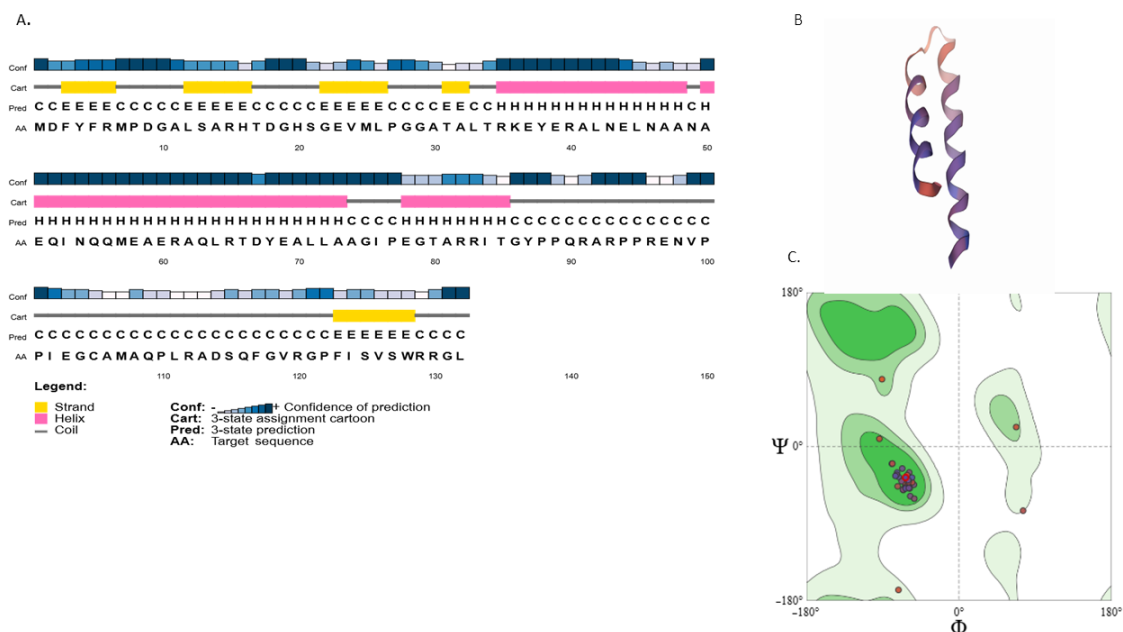


Figure 4. Structural analysis of HP2. Amino acid sequence of HP2, along with the structural elements, are shown in (A). The 3D structure created by Swiss-Model and the Ramachandran plot of HP2 are presented in (B,C), respectively. The conformationally most favourable regions are presented as red dots within the green area.

The failure of functional annotation of the HPs necessitated further attempts to examine if any bioactivity could be attributed to them by employing other available tools. In this endeavour, we applied the tools meant for the assessment of antimicrobial potential. All the HP sequences were subjected to assessment of antimicrobial potential (AMP) using a multi-model in silico pipeline (Table S6). Interestingly, HP2 showed strong evidence for AMP, with consistent results from three models. HP2 was predicted to be an antimicrobial peptide with a probability of >0.5. The DBAASP database revealed that it pertained to antiviral activity. Antibacterial or antifungal activities were not predicted. The identified antiviral targets included Hepatitis C virus (HCV), Zika virus (ZIKV), SARS-CoV, Japanese encephalitis virus (JEV), HIV-1, and DENV (Table 3). The target viral proteins were extracted from UniProt. To further validate the interaction between targets and HP2, molecular docking was performed using the HDock server (Figure 5). The peptide was able to bind with relatively high binding energies below -165 kcal/mol (Table 3), consistent with a receptor-ligand complex having the strongest binding affinity. HP2 demonstrated the lowest docking score, suggesting strong interaction potential for HIV and SARS-CoV-2, followed by others. JEV demonstrated the highest interface area and residue count, with the highest number of non-bonded contacts, suggesting dense molecular proximity and extensive surface contact with its peptide partner (Figure S6A). HIV and SARS-CoV-2 also appeared to form large interfaces, while DENZ2 and HCV showed smaller contact regions. HIV formed the most hydrogen bonds, implying strong specific interactions. DENZ2 and SARS-CoV-2 showed higher salt-bridge formation than others. HIV achieved the best confidence score and RMSD, indicating precise and reliable binding. SARS-CoV-2 also demonstrated strong performance. The visual of multi-metric docking efficiency in a compact circular form reveals overall trends effectively (Figure S6B). HIV had the highest RMSD, suggesting structural flexibility or docking instability. Conversely, SARS-CoV-2

and JEV demonstrated lower RMSD values, indicating tighter and more consistent peptide binding (Figure S6C). HIV and SARS-CoV-2 demonstrated superior binding confidence, aligning with biological relevance and stable complex formation. DENZ2, HCV, and JEV rank slightly lower, though still within acceptable confidence thresholds (Figure S6D). HIV and SARS-CoV-2 have the most favourable (lowest) docking energies, suggesting a strong interaction potential for antiviral activity. All values fell within biologically significant ranges, yet SARS-CoV-2's lower score appeared to enhance its binding reliability compared to HCV or DENZ2 (Figure S6E). Visual analyses such as heatmaps, radar plots, and pair plots revealed meaningful trends, such as the correlation between interface area and interaction density, which supports the interpretability of the docking outcomes.

Table 3. AMP prediction using Database of Antimicrobial Activity and Structure of Peptides of hypothetical proteins with moderate and strong AMP.

Based on DBAASP AVP Activity Prediction		HP2	
	Name	Class	Predictive value
Viral pathogen	Hepatitis C virus (HCV)	Active	0.8
	Zika virus (ZIKV)	Active	0.53
	West Nile virus (WNV)	Not Active	0.54
	SARS-CoV-2	Active	0.64
	SARS-CoV	Active	0.65
	Japanese encephalitis virus (JEV)	Active	0.74
	HIV-1	Active	0.53
	DENV-2	Active	0.55
	DENV-1	Active	0.6
Fungal pathogen	<i>Candida albicans</i>	Not active	-
	<i>Saccharomyces cerevisiae</i>		
Bacterial pathogen	Gram-positive Gram-negative	Not active	-
Hemolytic activity prediction of peptide: Human erythrocytes		Not active	-

For the development of a drug to target a protein, it is important to find the active site of the protein. One active site was identified by using the CASTp server. This site had 38 amino acids and a solvent-accessible (SA) surface area of 0.376. The key active residues of the pocket are THR³⁴, ARG³⁵, LYS³⁶, GLU³⁷, TYR³⁸, GLU³⁹, ARG⁴⁰, ALA⁴¹, LEU⁴², ASN⁴³, GLU⁴⁴, LEU⁴⁵, ASN⁴⁶, ALA⁴⁷, ALA⁴⁸, ASN⁴⁹, ALA⁵⁰, GLU⁵¹, GLN⁵², ILE⁵³, ASN⁵⁴, GLN⁵⁵, GLN⁵⁶, MET⁵⁷, GLU⁵⁸, ALA⁵⁹, GLU⁶⁰, ARG⁶¹, ALA⁶², GLN⁶³, LEU⁶⁴, ARG⁶⁵, THR⁶⁶, ASP⁶⁷, TYR⁶⁸, GLU⁶⁹, ALA⁷⁰, LEU⁷¹, and LEU⁷². The canonical SMILES of the HP2-active peptide revealed its physicochemical properties, supporting its potential as a pharmaceutical candidate. The predicted toxicological properties are presented in Table 4. The findings revealed that HP2 was similar to some of the known antiviral peptides (Table 4). These findings suggested HP2 as not an inhibitor of the human ether-a-go-go-related gene (hERG) that exhibits low rat toxicity, with a median lethal dose (LD₅₀) and no AMES toxicity. However, ADMET predictions suggest the acute oral toxicity of HP2, with an estimated LD₅₀ of <50 mg/kg.

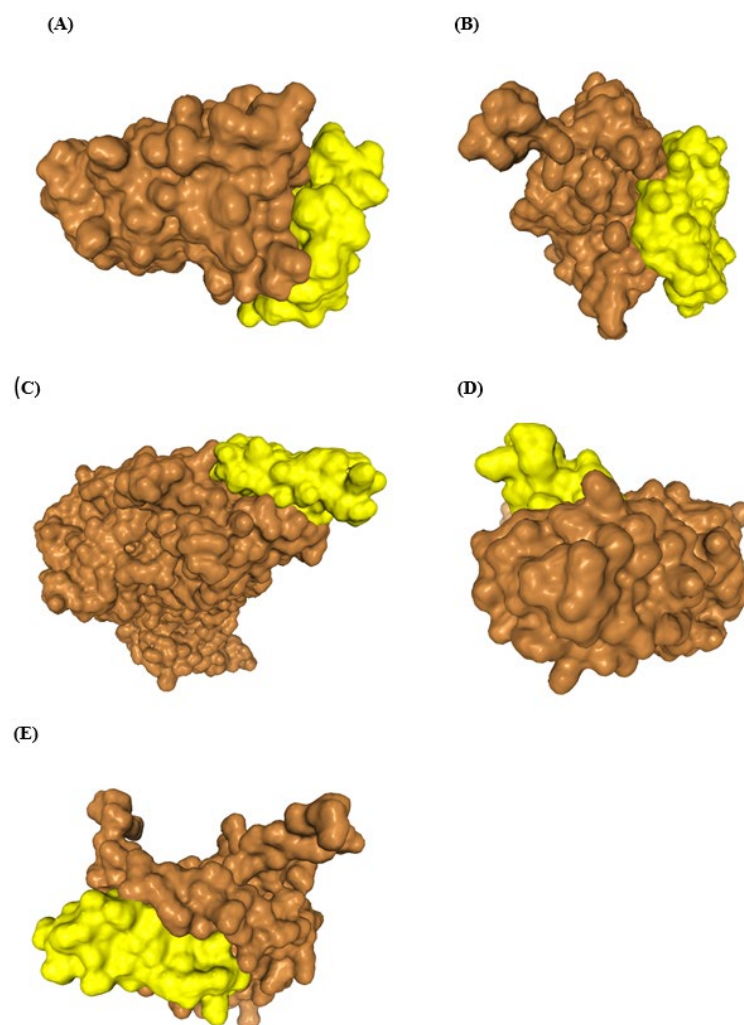


Figure 5. Molecular docking analysis of HP2 with viral proteins. The figure illustrates the predicted binding poses of HP2 with five viral proteins, highlighting key protein–peptide interactions. Panels (A–E) show surface representations of the complexes, with interacting residues indicated in brown and yellow. Specifically, (A) shows the DENV-HP2 complex, (B) HCV-HP2, (C) HIV-HP2, (D) JEV-HP2, (E) SARS-CoV-2-HP2, demonstrating the potential binding modes and interaction patterns of HP2 across diverse viral targets.

Table 4. Comparative pharmacokinetic properties of well-known AVP families isolated from different natural sources with HP2-active peptide (HP2-AP).

AVP		Enfuvirtide	C34 Peptide	hBD-1	Tifuvirtide	HP2-AP
Size of peptide	In amino acids	36	34	36	39	31
Type of peptide	Synthetic/natural	Natural	Synthetic	Natural	Synthetic	Natural
Physiochemical properties	Molecular weight (MW)	4448.16	4182.11	3931.78	4992.48	3558.74
Toxicity	hERG blocker	No	No	No	No	No
	AMES toxicity	No	No	No	No	No
	Rat oral acute toxicity	No	No	No	No	No

Table 4. Cont.

	AVP	Enfuvirtide	C34 Peptide	hBD-1	Tifuvirtide	HP2-AP
	Carcinogenicity	No	No	No	No	No
	LC50DM	5.392	5.238	6.096		4.987
Lipophilicity	Log Po/w	−0.919	−1.382	−1.643	−0.544	−4.55
Water solubility	Log S (ESOL)	−3.814	−4.02	−3.06	−3.785	−3.535
Pharmacokinetics	GI absorption	Low	Low	Low	Low	Low
Drug likeliness	Lipinski	Rejected	Rejected	Rejected	Rejected	Rejected
	Pfizer rule	Accepted	Accepted	Accepted	Accepted	Accepted
Excretion	T1/2	2.841	2.597	4.176	2.994	2.367
	CL _{plasma}	−0.25	−0.061	0.152	−0.475	0.179
	Accession no. (Drug bank/Uniprot)	DB00109 (FDA-approved)	P19550 (628–661 residue)	P60022	DB05413 (under trial)	-

4. Discussion

The study presented a detailed molecular characterization of a pharmaceutically significant actinobacterium, *S. chrestomyceticus* strain ADP4. Phylogenetic analysis provided an evolutionary insight among species, demonstrating ADP4 as a new strain of *S. chrestomyceticus* that was closest to the strain NPDC045757. The results suggested strain-level divergence within the strains of *S. chrestomyceticus*. Genomic attributes of all the strains stood in good agreement with one another. Parameters such as large genome size (9.31–9.86 Mbp), GC content (72.05 +/− 0.05%), number of coding sequences (8.3K + 0.3), and distribution of functional gene categories showed only minor variations, which suggested that the strains shared a common evolutionary background. The consistency in these features also ruled out the presence of anomalous genomes, thereby strengthening the reliability of comparative analyses. The genome size of the genus *Streptomyces* ranged from 4.8 Mbp to 13.6 Mbp, with a median of 8.5 Mbps [35]. The strains of *S. chrestomyceticus* possessed genomes that fell on the higher end of the size spectrum. The larger genome size might also involve the accumulation of non-coding DNA sequences or hypothetical ORFs [36]. Non-coding and coding contents of the genome of *S. chrestomyceticus* strains were noted to be 12.4–16.7% and 83.3–87.6%, respectively. For *Streptomyces* spp., higher coding DNA fractions (83% to 90%) were observed [37]. All the strains of *S. chrestomyceticus* corresponded to a slightly higher proportion of non-coding sequences, which may include regulatory elements or intergenic regions, suggesting potential specific genomic features. The number of CRISPRs varied in different organisms, and even differed among the strains of the same species [38]. For most of the *S. chrestomyceticus* strains, the number was either six or less. Strains ADP4 and NPDC003050 stood out with a comparatively very high number, thirteen. The type of CRISPR that is common in prokaryotes, including *Streptomyces*, is Type I-E. Type I-E is also found in all the *S. chrestomyceticus* strains. These natural CRISPR sequences of *Streptomyces* mostly function in protecting against phages and plasmids [39,40]. For strain ADP4, the presence of Type V-F2 CRISPR in addition to Type I-E, made it more distinct. This type certainly occurred in some bacteria with important applications in genome editing that could be harnessed through engineering efforts [41]. The large size, type/number of CRISPRs, strain-specific coding genes, duplications, and transposable elements comprised the discrete genomic features of the strain ADP4, which might contribute to strain complexity and specificity. The presence of unannotated entries indicated gaps in the existing database, which resulted in under-representation of the genome or-

ganization among prokaryotes [42]. Pathways for the biosynthesis of an extensive range of antimicrobial metabolites were found in genome-based predictions, suggesting strain ADP4 as an excellent source for the production of therapeutically significant metabolites. This included biosynthesis of various classes of secondary metabolites such as terpenoids, T1PKS/T2PKS, alkaloids, and others. The metabolic pathway showed the co-occurrence of resistance genes and genes involved in the biosynthesis of antibiotics, suggesting that the strain ADP4 possessed the self-resistance strategy [43]. The presence of strain-specific metabolic genes highlighted intra-species genomic diversity. The categories included glycan biosynthesis and metabolism, biosynthesis of other secondary metabolites, replication and repair, membrane transport, lipid biosynthesis proteins, and uncharacterized genes.

Among all the *Streptomyces* genomes, BGCs per genome ranged from 8 (*S. gilvigriseus* MUSC 26) to 83 (*S. rhizosphaericus* NRRL B-24304) [44]. The size of BGCs in ADP4 ranged from 4.9 kb to 90 kb. *Streptomyces* are known to have BGCs of varying sizes, ranging from 4.9 kb to 148 kb. While it was suggested that there could be a relation between genome size and the number of BGCs per genome [45], this was not evident among the strains of *S. chrestomyceticus*, where the number of BGCs varied significantly, but no corresponding variation within the size of the genomes was noted. *S. chrestomyceticus* is a promising species among *Streptomyces* owing to its repertoire of BGCs. They are known to produce compounds such as paramomycin, neomycin E/F, aminocidin, and pyrrolisatin, which are known to be produced by other *Streptomyces* [46]. Therapeutic potential of the strain ADP4 was evident from the analysis of its BGCs (Table 3). Gene clusters of metabolites ubiquitous to the genus *Streptomyces*, for example, ecotine, geosmin, and desferrioxamine, were found in the strain ADP4.

The genome mining for BGCs demonstrated the potential of strain ADP4 as a promising source of therapeutically significant metabolites (Table 3). The gene clusters of the metabolites ubiquitous to the genus *Streptomyces*, for example, ecotine, geosmin, and desferrioxamine, were found in strain ADP4. The presence of unique clusters (not predicted in other *S. chrestomyceticus*) such as lysolipin, notonesomycin A, and α naphthocyclinoic acid/fogacin suggested acquisition of the genes through horizontal gene transfer or lineage-specific gene gain events, potentially contributing to its unique metabolic or adaptive capabilities [47]. Moreover, the variability in BGCs as displayed in genome mining across *Streptomyces* spp. highlighted that species-level identity alone was insufficient for metabolic redundancy [48]. Clusters like antipain (NRPS), legonoxamine/desferrioxamine (siderophore), and purincyclamide (cyclodipeptide) were predicted with 100% similarity for all *S. chrestomyceticus* strains. Based on the evidence reported, purincyclamide appears as an *S. chrestomyceticus*-specific metabolite. The purincyclamide BGC is a 3.66 kb cluster reported first in *S. chrestomyceticus* with five genes (pcmA-E). The gene pcmA is an essential gene that acts as cyclodipeptide synthase and is present in all the strains of *S. chrestomyceticus*. The total size of the cluster here is 20.77 kb. BGCs exhibited complete (100%) identity, underscoring their conserved functional potential (Figure S7). These findings suggested that purincyclamide could act as a marker BGC for the identification of *S. chrestomyceticus*.

Based on the metabolic profile of ADP4, a unique metabolic protein, pks5 (KEGG ID: K12433), was found to be involved in T1PKS biosynthesis. This was observed to be unique to the strain ADP4, as pks5 was not present in other strains of *S. chrestomyceticus*. Further analysis revealed a 0.67 identity score of T1PKS BGC with BGC0001244.3. This cluster and other related clusters are known to produce coumarin and chromene classes of compounds. Recently, such a class of compounds was reported from strain ADP4 [7,8]. Thus, the metabolic protein pks5 is another strain-specific attribute of strain ADP4, conferring diversity to its antimicrobial profile.

Comparative genomics offered a comprehensive approach for a systematic analysis to understand the similarities and differences among the genomes from different strains of the same species. Comparative analyses of CDSs of strain ADP4 against other strains of *S. chrestomyceticus* were initially performed to evaluate strain-level specificity. These analyses revealed a set of genes unique to strain ADP4, including both accessory genes and core metabolic genes. To find novel strain-specific genes, protein sequences from them were subjected to BLAST analysis. Protein sequences reflect functional uniqueness more accurately as they are conserved across strains and reduce the confounding effects of synonymous mutations and codon bias more than DNA sequences. This analysis indicated that strain ADP4 possessed distinct genetic features, which were not observed in other *Streptomyces* spp. and might contribute to its specialized metabolic and biosynthetic potential. The presence of these strain-specific sequences underscored the importance of examining genomic variation at the strain level, rather than relying solely on species-level classification, to identify candidates for further functional investigation.

Interestingly, the identified strain-specific sequences in ADP4 matched with hypothetical proteins (HPs). Due to being uncharacterized proteins and the limitation of database knowledge, complete functional annotation for the HPs was a challenge. Comparative orthologous gene cluster analysis among the strains of *S. chrestomyceticus* revealed 154 singletons of ADP4. Further, eight strain-specific unique sequences were extracted. These sequences were HPs lacking characterized functions in the existing database. The identification of these HPs included targeted functional studies, including protein structure prediction, conserved domain analysis, and elucidating their biological roles. Functional characterization of the HPs by employing various platforms suggested that only two out of the eight, HP1 and HP2, could be annotated. First, HP1 appeared as an extracellular polypeptide with an N-terminal signal peptide domain. It had no detectable homology in the current databases, including no BLAST hit, suggesting it to be a truly unique protein. The lack of characterization and homology suggested that it might contribute to a unique pathway or adaptive trait. HP2 was identified as the second sequence, with no significant BLAST matches to the existing databases, highlighting its uniqueness. Functional annotation using various bioinformatic tools failed to assign any known function. Regarding antimicrobial potential of HP1, a moderate score was observed due to the lack of consistent activity across all the predictive models. Further, DBAASP prediction showed no antagonistic activity for HP1, suggesting it not to have AMP. However, HP2 displayed consistent antagonistic activity with a strong score and specific antiviral potential. The combination of uniqueness and AMP indicated that HP2 constituted a novel antiviral peptide specific to the strain ADP4. As unravelled by ProtParam analysis, the complete peptide of HP2 consisted of 132 aa, which appeared unstable, whereas the active peptide domain of 26 aa was predicted to be stable. The majority of AVPs have a peptide length of 11–30 amino acids. However, the antiviral potential of peptides depends largely on their amino acid composition rather than size [49]. Lysine has been reported as a key residue for AMPs [50]. Hydrophobic residues like Leu and Ile are key for differentiating AVPs from non-AVPs. AVPs have dominant hydrophobic residues, comprising >40% of total amino acids, with Leu and Arg the most preferred. Arg enhances the electrostatic binding to viral envelopes and Leu promotes hydrophobic interaction and helix stability, driving antiviral activity [51]. Structurally, helical regions are enriched in Lys and Leu, while beta-sheets of AVPs favour Arg and Cys; other AMP families are dominated by residues like Trp, Gly, His, Arg, or Pro, which critically determine their structure and activity. The structure of HP2 is likely helical because of stabilization by hydrophobic Leu and Ile residues, with multiple positively charged Arg and Lys residues facilitating electrostatic interactions. The presence of these

residues facilitates strong hydrophobic interactions, altogether driving its antiviral activity as described for typical AVPs.

The molecular docking studies of HP2 showed the most favourable interaction profile for HIV, with the lowest docking score (-221.55 kcal/mol), highest confidence score (0.8071), and highest ligand RMSD, suggesting strong but potentially flexible docking. Molecular docking revealed 19 residues of HIV forming bonds with 13 residues of HP2. These bonds include nine hydrogen bonds with Gln55, Ile53, Tyr68, Arg61, Asn46, Glu69, and Leu72, one salt bridge, and other non-bonded contacts. HP2 majorly bonded with the envelope glycoprotein gp120 variable region V2 (156–196). These variable loops with V2 are targeted by some inhibitors, which interfere with viral entry by blocking coreceptor binding or by inhibiting necessary conformational changes for fusion. SARS-CoV-2 exhibited a low RMSD (28.97 Å) and a high confidence score (0.7968), indicating a structurally stable and accurate binding orientation. JEV formed the most extensive interface, as evidenced by the highest interface area (1234 Å²) and the greatest number of non-bonded contacts (223), pointing to dense molecular interactions. DENZ2 and HCV showed moderate interaction profiles with less extensive interface residues and comparatively weaker docking metrics. Based on the results reported here, HP2 could be a novel antiviral peptide from the genus *Streptomyces*.

5. Conclusions

The study presented a comprehensive molecular characterization of the strain ADP4 from a genomic perspective, which led to the discovery of strain-specific molecular attributes. It has been demonstrated that strain ADP4 possessed a strain-specific metabolic pathway and BGCs; for example, pks5, associated with antimicrobial metabolites of coumarin and chromene classes. Also, eight ADP4-specific hypothetical proteins were discovered, which might play vital roles in the biology of the strain, that need to be understood through further studies. It was, however, unexpected to find broad-spectrum antiviral activity associated with one of the hypothetical proteins, HP2. The results on the antiviral potential of HP2 appear promising and need to be established through further in-depth studies.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/microbiolres16120249/s1>, Figure S1: Gene ontology (GO) Functional annotation of ADP4 genome. The clockwise direction represents the distribution of genes for different functional categories: biological processes, cellular processes, metabolic functions, and KEGG pathways. Figure S2: Pathway prediction. Figure S3: Structure of biosynthetic gene clusters. Figure S4: HP1 protein structural analysis. Figure S5: Multiple sequences alignment analysis. Figure S6: Molecular docking analysis. Figure S7: Synteny of BGC of Purincyclamide of *S. chrestomyceticus* strains. Table S1: List of genomes used for phylogenetic analysis along with their accession number, size and percent GC values. Table S2: COG analysis - number of genes associated with general cog functional categories. Table S3: List of putative BGCs predicted by antiSMASH 7.10 in the genome of *S. chrestomyceticus* ADP4 and their bioactivities. Table S4: COG analysis. Number of genes associated with general cog functional categories. Table S5: Physicochemical Properties of the HP1 and HP2 as predicted by the ProtParam tool. Table S6: Antimicrobial potential (AMP) assessment using a multi-model in silico pipeline of all-strain-specific hypothetical sequences.

Author Contributions: V.V. conducted experiments, analyzed data, and prepared the first draft of the manuscript. M.K.M. supervised the bioinformatics part of the study. A.K.D. conceptualized the study, acquired funding, and supervised the study. M.K.M. and A.K.D. reviewed and finalized the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The whole-genome sequence of the strain ADP4 has been deposited in GenBank (https://www.ncbi.nlm.nih.gov/datasets/genome/GCF_036562025.1/, accessed on 25 November 2025) under the BioProject accession PRJNA1007925, BioSample SAMN37105823, and GenBank assembly accession GCA_036562025.1. All the data pertaining to this work have been included in the manuscript and provided in the Supplementary File.

Acknowledgments: The authors acknowledge Netaji Subhas University of Technology, New Delhi, for providing the materials and administrative support to carry out the present study. Acknowledgement is also extended to the University Science Instrumentation Centre, University of Delhi, for the use of LCMS/MS facilities during the course of this investigation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Selim, M.S.M.; Abdelhamid, S.A.; Mohamed, S.S. Secondary metabolites and biodiversity of actinomycetes. *J. Genet. Eng. Biotechnol.* **2021**, *19*, 72. [CrossRef]
2. Kim, S.; Lim, S.W.; Choi, J. Drug discovery inspired by bioactive small molecules from nature. *Anim. Cells Syst.* **2022**, *26*, 254–265. [CrossRef] [PubMed]
3. Wang, S.; Li, W.; Wang, Z.; Yang, W.; Li, E.; Xia, X.; Yan, F.; Chiu, S. Emerging and reemerging infectious diseases: Global trends and new strategies for their prevention and control. *Signal Transduct. Target Ther.* **2024**, *9*, 223. [CrossRef]
4. Helmi, N.R. Exploring the diversity and antimicrobial potential of actinomycetes isolated from different environments in Saudi Arabia: A systematic review. *Front. Microbiol.* **2025**, *16*, 1568899. [CrossRef]
5. Srivastava, V.; Dubey, A.K. Anti-Biofilm Activity of the Metabolites of *Streptomyces chrestomyceticus* Strain ADP4 against *Candida albicans*. *J. Biosci. Bioeng.* **2016**, *122*, 434–440. [CrossRef]
6. Singh, R.; Dubey, A.K. Differential Synthesis of Secondary Metabolites by *Streptomyces chrestomyceticus* Strain ADP4 in Response to Modulation in Nitrogen Source and Its Anti-*Candida* Activity. *Proceedings* **2020**, *66*, 5. [CrossRef]
7. Singh, R.; Ali, M.; Dubey, A.K. Anti-*Candida* Attributes and In-Silico Drug-Likeness Properties of Phenyl 2'β, 6'β-Trimethyl Cyclohexyl Ketone and Phenyl Nonanyl Ether Produced by *Streptomyces chrestomyceticus* ADP4. *J. Appl. Microbiol.* **2023**, *134*, lxac024. [CrossRef]
8. Singh, R.; Ali, M.; Dubey, A.K. Phenyl Pentyl Ketone and m-Isobutyl Methoxy Benzoate Produced by *Streptomyces chrestomyceticus* ADP4 Are Potent Antimicrobial Agents Displaying Broad Spectrum Activities. *Indian J. Microbiol.* **2023**, *63*, 181–189. [CrossRef]
9. Singh, R.; Ali, M.; Dubey, A.K. Identification and Characterization of a Novel Decalin Derivative with Anti-*Candida* Activity from *Streptomyces chrestomyceticus* Strain ADP4. *Arch. Microbiol.* **2024**, *206*, 50. [CrossRef] [PubMed]
10. Singh, R.; Shukla, J.; Ali, M.; Dubey, A.K. A Novel Diterpenic Derivative Produced by *Streptomyces chrestomyceticus* ADP4 Is a Potent Inhibitor of Biofilm and Virulence Factors in *Candida albicans* and *C. auris*. *J. Appl. Microbiol.* **2024**, *135*, lxae139. [CrossRef]
11. Singh, R.; Shukla, J.; Ali, M.; Dubey, A.K. A Novel Benzopyrone Derivative from *Streptomyces chrestomyceticus* ADP4 Inhibits Growth and Virulence Factors of *Candida albicans*. *Curr. Microbiol.* **2025**, *82*, 201. [CrossRef]
12. Srivastava, V.; Singla, R.K.; Dubey, A.K. Inhibition of Biofilm and Virulence Factors of *Candida albicans* by Partially Purified Secondary Metabolites of *Streptomyces chrestomyceticus* Strain ADP4. *Curr. Top. Med. Chem.* **2018**, *18*, 925–945. [CrossRef]
13. Gandham, P.; Vadla, N.; Saji, A.; Srinivas, V.; Ruperao, P.; Selvanayagam, S.; Saxena, R.K.; Rathore, A.; Gopalakrishnan, S. Genome assembly, comparative genomics, and identification of genes/pathways underlying plant growth-promoting traits of an actinobacterial strain, *Amycolatopsis* sp. (BCA-696). *Sci. Rep.* **2024**, *14*, 15934. [CrossRef] [PubMed]
14. Abraha, H.B.; Ramesha, R.M.; Ferdiansyah, M.K.; Son, H.; Kim, G.; Park, B.; Jeong, D.Y.; Kim, K.P. Genome Analysis of a Newly Sequenced *B. subtilis* SRCM117797 and Multiple Public *B. subtilis* Genomes Unveils Insights into Strain Diversification and Biased Core Gene Distribution. *Curr. Microbiol.* **2024**, *81*, 305. [CrossRef]
15. Otani, H.; Uduary, D.W.; Mouncey, N.J. Comparative and pangenomic analysis of the genus *Streptomyces*. *Sci. Rep.* **2022**, *12*, 18909. [CrossRef]
16. Minas, K.; McEwan, N.R.; Newbold, C.J.; Scott, K.P. Optimization of a high-throughput CTAB-based protocol for the extraction of qPCR-grade DNA from rumen fluid, plant and bacterial pure cultures. *FEMS Microbiol. Lett.* **2011**, *325*, 162–169. [CrossRef]
17. Andrews, S. FastQC: A Quality Control Tool for High Throughput Sequence Data. 2010. Available online: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (accessed on 31 March 2021).
18. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **2014**, *30*, 2114–2120. [CrossRef]

19. Bankevich, A.; Nurk, S.; Antipov, D.; Gurevich, A.A.; Dvorkin, M.; Kulikov, A.S.; Lesin, V.M.; Nikolenko, S.I.; Pham, S.; Pribelski, A.D.; et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* **2012**, *19*, 455–477. [[CrossRef](#)] [[PubMed](#)]
20. Parks, D.H.; Imelfort, M.; Skennerton, C.T.; Hugenholtz, P.; Tyson, G.W. CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **2015**, *25*, 1043–1055. [[CrossRef](#)] [[PubMed](#)]
21. Simão, F.A.; Waterhouse, R.M.; Ioannidis, P.; Kriventseva, E.V.; Zdobnov, E.M. BUSCO: Assessing Genome Assembly and Annotation Completeness with Single-Copy Orthologs. *Bioinformatics* **2015**, *31*, 3210–3212. [[CrossRef](#)]
22. Tanizawa, Y.; Fujisawa, T.; Nakamura, Y. DFAST: A flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **2018**, *34*, 1037–1039. [[CrossRef](#)] [[PubMed](#)]
23. Haft, D.H.; Badretdin, A.; Coulouris, G.; DiCuccio, M.; Durkin, A.S.; Jovenitti, E.; Li, W.; Mersha, M.; O'Neill, K.R.; Virothaisakun, J.; et al. RefSeq and the prokaryotic genome annotation pipeline in the age of metagenomes. *Nucleic Acids Res.* **2024**, *52*, D762–D769. [[CrossRef](#)]
24. Chaudhari, N.M.; Gupta, V.K.; Dutta, C. BPGA—An ultra-fast pan-genome analysis pipeline. *Sci. Rep.* **2016**, *6*, 24373. [[CrossRef](#)]
25. Tatusov, R.L.; Galperin, M.Y.; Natale, D.A.; Koonin, E.V. The COG database: A tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **2000**, *28*, 33–36. [[CrossRef](#)]
26. Kanehisa, M.; Goto, S.; Kawashima, S.; Okuno, Y.; Hattori, M. The KEGG resource for deciphering the genome. *Nucleic Acids Res.* **2004**, *32*, D277–D280. [[CrossRef](#)]
27. Bortolaia, V.; Kaas, R.S.; Ruppe, E.; Roberts, M.C.; Schwarz, S.; Cattoir, V.; Philippon, A.; Allesoe, R.L.; Rebelo, A.R.; Florensa, A.F.; et al. ResFinder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* **2020**, *75*, 3491–3500. [[CrossRef](#)]
28. Alcock, B.P.; Raphenya, A.R.; Lau, T.T.Y.; Tsang, K.K.; Boucharde, M.; Edalatmand, A.; Huynh, W.; Nguyen, A.-L.V.; Cheng, A.A.; Liu, S.; et al. CARD 2020: Antibiotic resistance surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res.* **2020**, *48*, D517–D525. [[CrossRef](#)]
29. Blin, K.; Shaw, S.; Augustijn, H.E.; Reitz, Z.L.; Biermann, F.; Alanjary, M.; Fetter, A.; Terlouw, B.R.; Metcalf, W.W.; Helfrich, E.J.N.; et al. antiSMASH 7.0: New and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* **2023**, *51*, W46–W50. [[CrossRef](#)] [[PubMed](#)]
30. Wang, J.; Chitsaz, F.; Derbyshire, M.K.; Gonzales, N.R.; Gwadz, M.; Lu, S.; Marchler, G.H.; Song, J.S.; Thanki, N.; Yamashita, R.A.; et al. The conserved domain database in 2023. *Nucleic Acids Res.* **2023**, *51*, D384–D388. [[CrossRef](#)]
31. Yoon, S.H.; Ha, S.M.; Lim, J.; Kwon, S.; Chun, J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Leeuwenhoek* **2017**, *110*, 1281–1286. [[CrossRef](#)] [[PubMed](#)]
32. Meier-Kolthoff, J.P.; Carbasse, J.S.; Peinado-Olarte, R.L.; Göker, M. TYGS and LPSN: A database tandem for fast and reliable genome-based classification and nomenclature of prokaryotes. *Nucleic Acids Res.* **2022**, *50*, D801–D807. [[CrossRef](#)] [[PubMed](#)]
33. Letunic, I.; Bork, P. Interactive Tree Of Life (iTOL) v5: An online tool for phylogenetic tree display and annotation. *Nucleic Acids Res.* **2021**, *49*, W293–W296. [[CrossRef](#)]
34. Sun, J.; Lu, F.; Luo, Y.; Bie, L.; Xu, L.; Wang, Y. OrthoVenn3: An integrated platform for exploring and visualizing orthologous data across genomes. *Nucleic Acids Res.* **2023**, *51*, W397–W403. [[CrossRef](#)]
35. Mohite, O.S.; Jørgensen, T.S.; Booth, T.J.; Charusanti, P.; Phaneuf, P.V.; Weber, T.; Palsson, B.O. Pangenome mining of the *Streptomyces* genus redefines species' biosynthetic potential. *Genome Biol.* **2025**, *26*, 9. [[CrossRef](#)] [[PubMed](#)]
36. Konstantinidis, K.T.; Tiedje, J.M. Trends between gene content and genome size in prokaryotic species with larger genomes. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 3160–3165. [[CrossRef](#)] [[PubMed](#)]
37. Tian, X.; Zhang, Z.; Yang, T.; Chen, M.; Li, J.; Chen, F.; Yang, J.; Li, W.; Zhang, B.; Zhang, Z.; et al. Comparative genomics analysis of *Streptomyces* species reveals their adaptation to the marine environment and their diversity at the genomic level. *Front. Microbiol.* **2016**, *7*, 998. [[CrossRef](#)]
38. Ishino, Y.; Krupovic, M.; Forterre, P. History of CRISPR-Cas from Encounter with a Mysterious Repeated Sequence to Genome Editing Technology. *J. Bacteriol.* **2018**, *200*, e00580-17. [[CrossRef](#)]
39. Loureiro, A.; da Silva, G.J. CRISPR-Cas: Converting a Bacterial Defence Mechanism into a State-of-the-Art Genetic Manipulation Tool. *Antibiotics* **2019**, *8*, 18. [[CrossRef](#)]
40. Qiu, Y.; Wang, S.; Chen, Z.; Guo, Y.; Song, Y. An Active Type I-E CRISPR-Cas System Identified in *Streptomyces avermitilis*. *PLoS ONE* **2016**, *11*, e0149533. [[CrossRef](#)]
41. Hua, H.M.; Xu, J.F.; Huang, X.S.; Zimin, A.A.; Wang, W.F.; Lu, Y.H. Low-Toxicity and High-Efficiency *Streptomyces* Genome Editing Tool Based on the Miniature Type V-F CRISPR/Cas Nuclease AsCas12f1. *J. Agric. Food Chem.* **2024**, *72*, 5358–5367. [[CrossRef](#)]
42. Lobb, B.; Tremblay, B.J.; Moreno-Hagelsieb, G.; Doxey, A.C. An assessment of genome annotation coverage across the bacterial tree of life. *Microb. Genom.* **2020**, *6*, e000341. [[CrossRef](#)] [[PubMed](#)]
43. Chen, X.; Pan, H.-X.; Tang, G.-L. Newly Discovered Mechanisms of Antibiotic Self-Resistance with Multiple Enzymes Acting at Different Locations and Stages. *Antibiotics* **2023**, *12*, 35. [[CrossRef](#)]

44. Belknap, K.C.; Park, C.J.; Barth, B.M.; Andam, C.P. Genome mining of biosynthetic and chemotherapeutic gene clusters in *Streptomyces* bacteria. *Sci. Rep.* **2020**, *10*, 2003. [[CrossRef](#)]
45. Shi, Y.M.; Hirschmann, M.; Shi, Y.N.; Ahmed, S.; Abebew, D.; Tobias, N.J.; Grün, P.; Cames, J.J.; Pöschel, L.; Kutenlochner, W.; et al. Global analysis of biosynthetic gene clusters reveals conserved and unique natural products in entomopathogenic nematode-symbiotic bacteria. *Nat. Chem.* **2022**, *14*, 701–712. [[CrossRef](#)]
46. Alam, K.; Mazumder, A.; Sikdar, S.; Zhao, Y.-M.; Hao, J.; Song, C.; Wang, Y.; Sarkar, R.; Islam, S.; Zhang, Y.; et al. *Streptomyces*: The biofactory of secondary metabolites. *Front. Microbiol.* **2022**, *13*, 968053. [[CrossRef](#)]
47. Vicente, C.M.; Thibessard, A.; Lorenzi, J.N.; Benhadj, M.; Hôtel, L.; Gacemi-Kirane, D.; Lespinet, O.; Leblond, P.; Aigle, B. Comparative genomics among closely related *Streptomyces* strains revealed specialized metabolite biosynthetic gene cluster diversity. *Antibiotics* **2018**, *7*, 86. [[CrossRef](#)]
48. Cimermancic, P.; Medema, M.H.; Claesen, J.; Kurita, K.; Wieland Brown, L.C.; Mavrommatis, K.; Pati, A.; Godfrey, P.A.; Koehrsen, M.; Clardy, J.; et al. Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **2014**, *158*, 412–421. [[CrossRef](#)] [[PubMed](#)]
49. Shah, J.N.; Guo, G.Q.; Krishnan, A.; Ramesh, M.; Katari, N.K.; Shahbaaz, M.; Abdellattif, M.H.; Singh, S.K.; Dua, K. Peptides-Based Therapeutics: Emerging Potential Therapeutic Agents for COVID-19. *Therapie* **2022**, *77*, 319–328. [[CrossRef](#)]
50. Neghabi Hajigha, M.; Hajikhani, B.; Vaezjalali, M.; Samadi Kafil, H.; Kazemzadeh Anari, R.; Goudarzi, M. Antiviral and antibacterial peptides: Mechanisms of action. *Heliyon* **2024**, *10*, e40121. [[CrossRef](#)] [[PubMed](#)]
51. Zakaryan, H.; Chilingaryan, G.; Arabyan, E.; Serobian, A.; Wang, G. Natural antimicrobial peptides as a source of new antiviral agents. *J. Gen. Virol.* **2021**, *102*, 001661. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.