



Review

Interaction of Proteins with Inverted Repeats and Cruciform Structures in Nucleic Acids

Richard P. Bowater ^{1,†} , Natálie Bohálová ^{2,3,†} and Václav Brázda ^{2,*}

¹ School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK; r.bowater@uea.ac.uk

² Department of Biophysical Chemistry and Molecular Oncology, Institute of Biophysics of the Czech Academy of Sciences, 61265 Brno, Czech Republic; nataliabohalova@gmail.com

³ Department of Experimental Biology, Faculty of Science, Masaryk University, Kamenice 5, 62500 Brno, Czech Republic

* Correspondence: vaclav@ibp.cz

† These authors contributed equally to this work.

Abstract: Cruciforms occur when inverted repeat sequences in double-stranded DNA adopt intra-strand hairpins on opposing strands. Biophysical and molecular studies of these structures confirm their characterization as four-way junctions and have demonstrated that several factors influence their stability, including overall chromatin structure and DNA supercoiling. Here, we review our understanding of processes that influence the formation and stability of cruciforms in genomes, covering the range of sequences shown to have biological significance. It is challenging to accurately sequence repetitive DNA sequences, but recent advances in sequencing methods have deepened understanding about the amounts of inverted repeats in genomes from all forms of life. We highlight that, in the majority of genomes, inverted repeats are present in higher numbers than is expected from a random occurrence. It is, therefore, becoming clear that inverted repeats play important roles in regulating many aspects of DNA metabolism, including replication, gene expression, and recombination. Cruciforms are targets for many architectural and regulatory proteins, including topoisomerases, p53, Rif1, and others. Notably, some of these proteins can induce the formation of cruciform structures when they bind to DNA. Inverted repeat sequences also influence the evolution of genomes, and growing evidence highlights their significance in several human diseases, suggesting that the inverted repeat sequences and/or DNA cruciforms could be useful therapeutic targets in some cases.

Keywords: cruciform; DNA base sequence; DNA structure; DNA supercoiling; epigenetics; genome stability; inverted repeat; replication; transcription



Citation: Bowater, R.P.; Bohálová, N.; Brázda, V. Interaction of Proteins with Inverted Repeats and Cruciform Structures in Nucleic Acids. *Int. J. Mol. Sci.* **2022**, *23*, 6171. <https://doi.org/10.3390/ijms23116171>

Academic Editor: Mateus Webba Da Silva

Received: 29 April 2022

Accepted: 30 May 2022

Published: 31 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

The wealth of DNA sequence information provided by genome sequencing projects has brought new insights into the primary sequences of genomes and also about possible sequence-dependent local secondary structures [1]. The primary base sequence alone is insufficient to decipher all principles that support basic molecular processes and those that maintain genomic and cellular stability. Inevitably, in-depth knowledge of epigenetic modifications and the local and global DNA structure is crucial for a full understanding of these processes. DNA molecules typically form two-stranded, right-handed helical B-form structures, which maximize the thermodynamic stability of the molecule [2]. However, a range of alternative (non-B) structures can also occur in DNA, and these are usually characterized by the occurrence of single-stranded regions (loops) and/or sites of disrupted base pair stacking (junctions between continuous B-form DNA and the alternative structure) [3]. Any disruption of stacking interactions or hydrogen bonds in base pairs alters

the thermodynamic stability of the molecule, but non-B DNA structures can be favourable for some sequences under some environmental (and cellular) conditions. Although they were initially considered as *in vitro* artefacts, several local secondary DNA structures are now well characterized and confirmed to form in living cells under physiologically relevant conditions [4–6]. These sequence-dependent conformational changes give rise to triplexes [7,8], G-quadruplexes [5,9], i-motifs [10], R-loops [8,11], four-way junctions [12], and cruciforms [13–15]. The latter is formed in DNA molecules containing inverted repeat sequences, either uninterrupted or interspaced with several additional bases forming loops. Thus, cruciform structures consist of branch-points, stems, and loops (Figure 1A) [15]. The thermodynamic stability of cruciforms is influenced by their size, with stable cruciforms usually requiring the inverted repeat to be at least six bases in length (for the stem, or one half of the repeat). Cruciforms can also arise from imperfect inverted repeats, meaning that unpaired bases occur within the stems of the cruciform, although this means such structures are energetically disfavoured compared to the fully base-paired structure [15,16]. In addition to inverted repeat unit size and unpaired bases, the length of the loop is also a critical factor influencing the stability of such structures (Figure 1B). Analyses of inverted repeats in various genomes have shown they have a non-random distribution and a functional association with regulatory sites, including promoters [17,18].

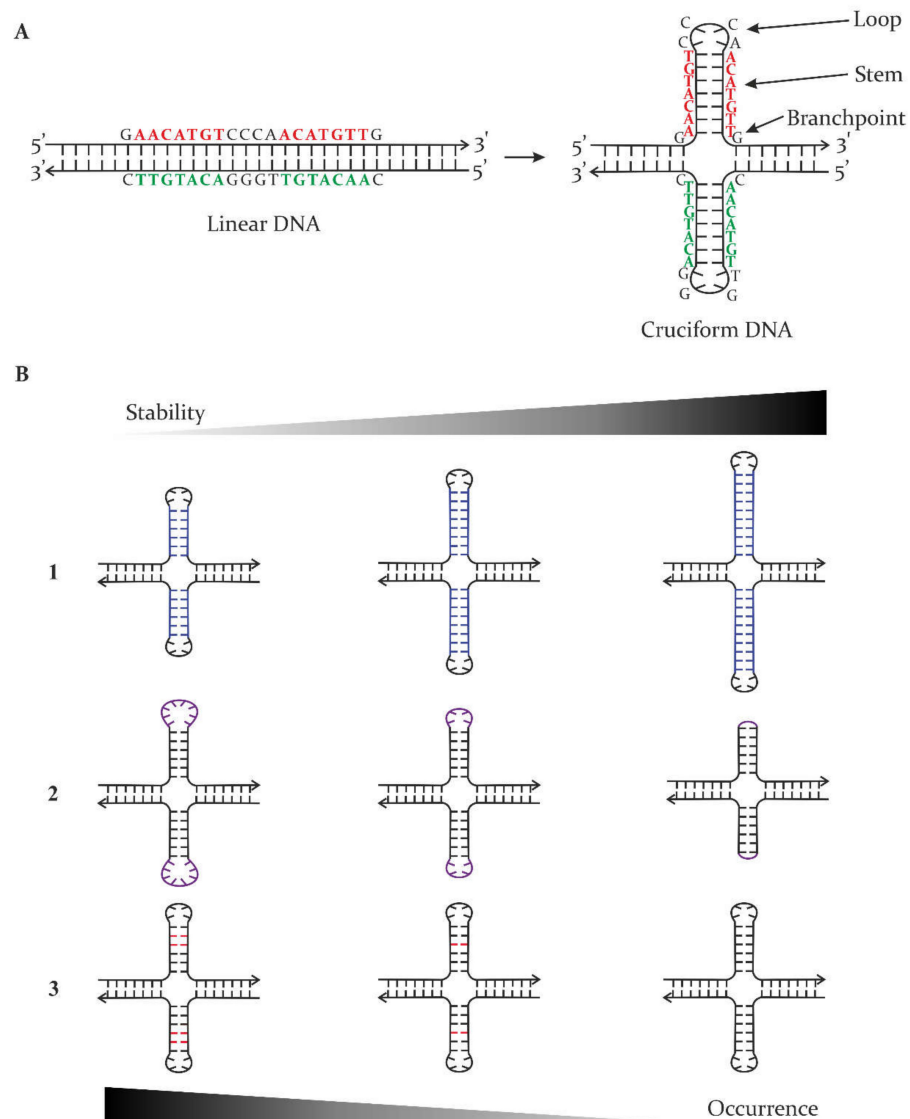


Figure 1. Inverted repeat sequences can form different types of double-stranded conformations.

(A) Transition of inverted repeat in a linear conformation to a double hairpin, cruciform state. For the sequence indicated, the cruciform structure consists of four branchpoints and two 7 bp-long stems, each with 4 nt loops. (B) Decisive factors for the resulting thermodynamic stability and genomic occurrence of cruciform structures are: (1) stem size indicated in blue; (2) loop length indicated in purple; (3) possible mismatches in base pairing indicated in red. The arrows at the top and bottom of part (B) highlight the relative stability and occurrence of the represented cruciforms, with the larger and darker part of the arrows indicating those that are most stable and are most likely to occur in genomes. For all schematic molecules, the arrow indicates the 3'-end of the DNA strand.

Inverted repeats and cruciforms have been found in all forms of life and appear to share similar functions and properties in many of them [3,6,17–21]. Inverted repeats are found in bacteria, eukaryotes, archaea, and viruses in higher amounts than would be expected from a random distribution of bases in both coding and non-coding regions, with a more pronounced frequency in non-coding regions. The frequency of inverted repeats in all organisms decreases with increasing length, but in most cases, the relative difference between expected and actual numbers tends to be higher for longer repeats [18]. As we describe in detail below, in all organisms, the presence of inverted repeats contributes to reduced genomic stability, primarily through the induction of inversions and the formation of hairpins and four-way junctions, which induce the stalling of polymerases and the generation of double-strand breaks. Cruciform conservation across all domains is, thus, a likely result of their involvement in essential molecular processes, such as opening of the DNA double helix during replication, transcription, and DNA damage repair [15].

2. Biophysical and Molecular Characterization of Cruciforms

The formation of cruciforms during the expression of genes was first postulated more than 50 years ago [22]. Their presence and function was subsequently studied both in vitro and in vivo, mostly for those located in plasmid DNAs from bacteria and yeasts [15]. The formation of cruciform structures requires the double-stranded helix of DNA to be opened, an energetically unfavourable process. A wide range of chemical and molecular probes have characterized properties that influence this process [6,23], with computer modelling methods helping to interpret experimental data [24]. Biophysical and molecular studies have clearly demonstrated that cruciforms are stable for some DNA molecules in vitro, but the situation has been less clear in vivo, mainly due to difficulties with studying the DNA structure inside cells. To assay for cruciform structures in cells, a range of probes of DNA structure have been used, including various factors that attack single-stranded regions of DNA, including psoralen and UV light cross-linking [6,25,26]. In some cases, the experiments cause the death of the cells, leading to studies being referred to as in situ to highlight that the cells are under physiological conditions, but may no longer be “living” [27]. Using *Escherichia coli* as a model, experiments have shown that large inverted repeats can be detected in cruciforms under some conditions, but sometimes at relatively low proportions of the total DNA [6]. Direct visualization of cruciforms in cells was attempted with a monoclonal antibody (2D3) shown to recognize cruciforms, but not heteroduplex slipped-stranded DNA containing a hairpin on one strand only [6,27]. Immunoprecipitation using this antibody revealed the presence of cruciform-containing DNA at a yeast replication origin, although it is unclear whether it specifically binds cruciforms or a panel of slipped-stranded DNA molecules [6,28]. Many methods continue to be used to study cruciform structures and their formation, from broad bioinformatic studies and electrophoretic in vitro assays to in vivo visualization by specific antibody interaction and single-molecule-level analyses [29–31]. Indeed, single-molecule manipulation of DNAs allowed cruciform formation, dynamics, and removal to be studied in real-time [32,33], as well as to reveal the mechanochemical properties of cruciform structure and cooperativity between opposing stem-loop structures [34].

In recent years, advances have been especially striking in high-resolution analyses of non-B DNA structures either as the nucleic acid alone or in combination with proteins.

In the context of this review, significant progress has been made in studies of four-way junctions, which are equivalent to the central part of cruciform structures—see Figure 1. Four-way junctions (often referred to as Holliday junctions) are critical intermediates in many DNA recombination and repair pathways [35], but it is important to recognize that such structures are usually formed by DNA molecules that do not contain inverted repeat sequences. A range of structural studies demonstrate that four-way junctions adopt different structures depending on the ionic environment and other factors [35,36]. X-ray crystallography and nuclear magnetic resonance (NMR) analyses of several DNA inverted repeat sequences confirm that they adopt the “stacked-X structure” in the absence of proteins, in which duplexes coaxially stack on each other. In thermodynamic terms, this type of structure has the most favourable energetics when monovalent or divalent cations are available to counteract the repelling interactions that occur between the negatively charged backbones, although cations are not an absolute requirement for the formation of stable cruciform structures. Figure 2 shows several views of a DNA inverted repeat structure determined at 2.10 Å for the sequence 5'-CCGGTACCGG-3' [37], and similar structures have been observed for a variety of other inverted repeats [36]. The DNA forms a four-way junction in a “stacked-X” conformation (Figure 2). Two strands are “continuous” and are closest to a B-DNA conformation, while the other two strands make a tight U-turn and cross at the junction. The stacked-X structure is seen clearly in Figure 2A,B. For this complex, a Na⁺ ion at its centre reduces electrostatic repulsion as the phosphodiester backbones come close to each other at the junction crossover (Figure 2C). Note that when the stacked-X structure is viewed from one face, Na⁺ is relatively protected by the DNA backbones, but it is relatively accessible to the local environment from the opposite side. Molecular dynamics simulation of a decamer inverted repeat as a four-way junction confirms its twofold symmetry and that temperature and its structural integrity are preserved by a range of other parameters (i.e., the presence of ions, solvents, etc.) [38]. Epigenetic markers on DNA, such as hydroxymethyl and methyl substituents, can be accommodated without disrupting the structure or stability of the cruciform, although they open the structure to make the junction core more accessible [36]. The binding of proteins—usually enzymes—to four-way junctions can alter their conformation, although they can have dramatically different effects [36,39–41]. High-resolution structures that are currently available for these altered conformations of four-way DNA junctions with proteins bound are usually for sequences that are not inverted repeats. It is expected that DNA cruciforms formed by inverted repeats will have similar flexibility when proteins bind to them, but this still has to be verified by high-resolution structures.

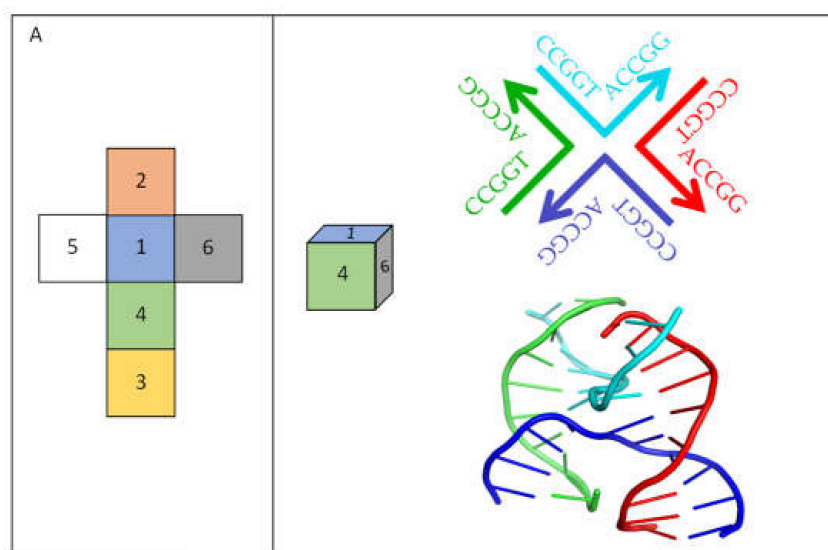


Figure 2. Cont.

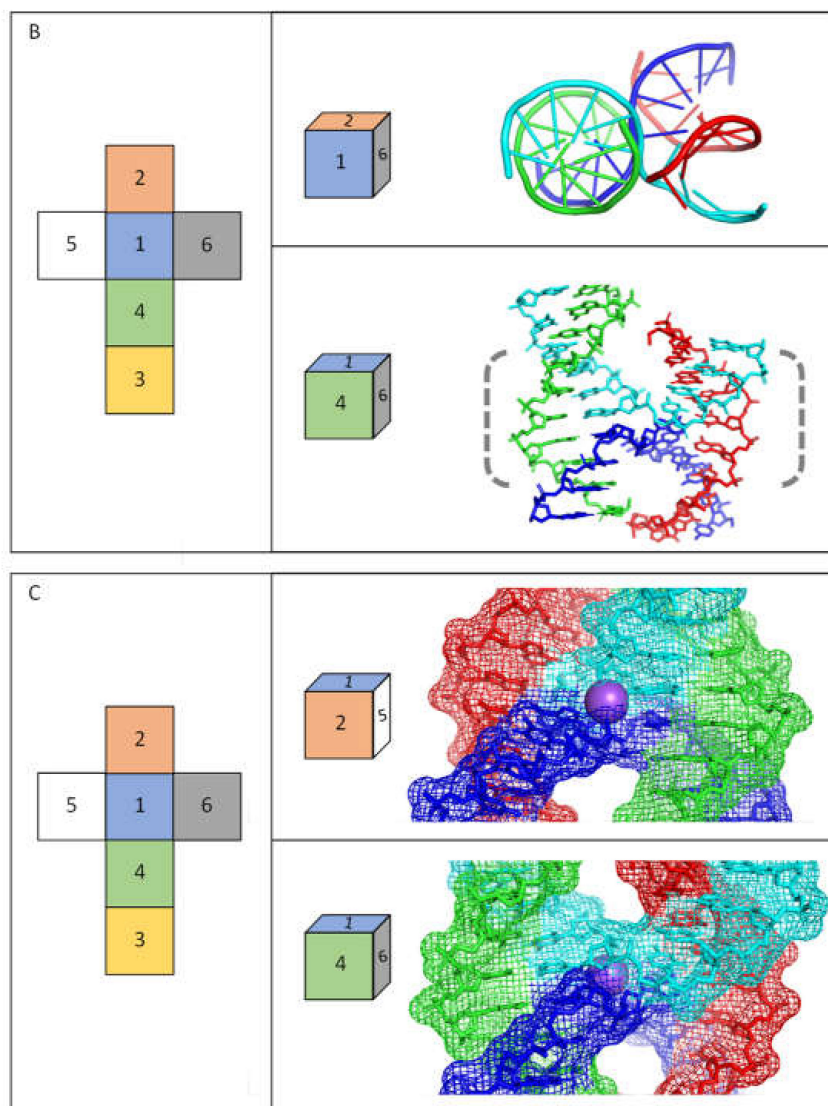


Figure 2. High-resolution structure of a cruciform (four-way junction) formed by an inverted repeat DNA sequence. Images show the X-ray crystallographic structure determined at 2.10 Å for DNA with the sequence 5'-CCGGTACCGG-3' (1DCW) [37]. The DNA alone forms a four-way junction in a stacked-X conformation, in which duplexes coaxially stack, with each pair of stacked duplexes related by +30° to +60° (right-handed) rotation. The continuous (least distorted relative to B-DNA) strands are coloured as green and red, while those of the crossing strands (making a tight U-turn) are coloured blue and cyan. In each panel, the images show the structure visualised via different axes viewpoints as indicated by the coloured squares. (A) The upper image provides a schematic view of the molecule, the distinct strands (in different colours), and their sequences, with arrows indicating the 3'-ends of the DNA strands. The lower image presents the high-resolution structure of 1DCW, illustrating its arrangement of base pairs. (B) The upper image views the structure down the helix axis of one pair of stacked duplexes, while the lower image views it from a rotational shift of approximately 90°. (C) The images zoom in on the central part of the structure (dashed bracketed region in (B)) to highlight the electrostatic interactions, particularly close to the Na⁺ ion at its centre. The lower image views the same face of the dyad axis shown in (B), and the upper image shows the opposite face of the axis, viewed from a rotational shift of approximately 180°.

3. Presence of Inverted Repeats in Genomes

The various experimental methods referred to above have provided abundant evidence for the presence of inverted repeats in genomes across all forms of life [6,42]. Since the start

of the 21st Century, the evidence has improved due to dramatic advances in sequencing technologies and bioinformatic analyses identifying genome sequences for many different organisms. Notably, it has been challenging to accurately sequence genomic regions that are rich in repeated bases for various reasons, but potentially including the presence of thermodynamically stable secondary structures [43]. Recent advances in sequencing methods mean that such problems can now usually be resolved, even for the human genome [44,45]. Here, we summarize the deepening understanding about the amounts of inverted repeats across all forms of life.

3.1. Viruses

Inverted repeats are found in higher numbers in many viral genomes than is expected from a random occurrence of bases [46]. This is true for many different types of viruses, but we illustrate this using Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) and adeno-associated viruses (AAV), which have single-stranded RNA and DNA genomes, respectively. Using the SARS-CoV-2 virus genome as an example, a total of 1203 inverted repeats with stems of 6–13 bp in length were identified. The average frequency of their occurrence was 40.24 inverted repeats per 1000 nt, whereas it was 33.90 for the entire *Nidovirales* family to which SARS-CoV-2 belongs [42]. Recurrent mutations were shown to occur within inverted repeats with a higher frequency than would be expected from a random distribution of them [47,48]. Furthermore, an abundance of inverted repeats was found within 5' untranslated regions of the *Nidovirales* family (Figure 3) [42]. In a different virus, AAV, terminal inverted repeats of 125 bases can form T-shaped hairpin structures by base-pairing of two small internal inverted repeat sequences and large flanking inverted repeat sequences [49]. This terminal inverted repeat of AAV was determined as the binding site for several transcriptional transactivators and was shown to facilitate recombination of the viral genome with the cellular genome.

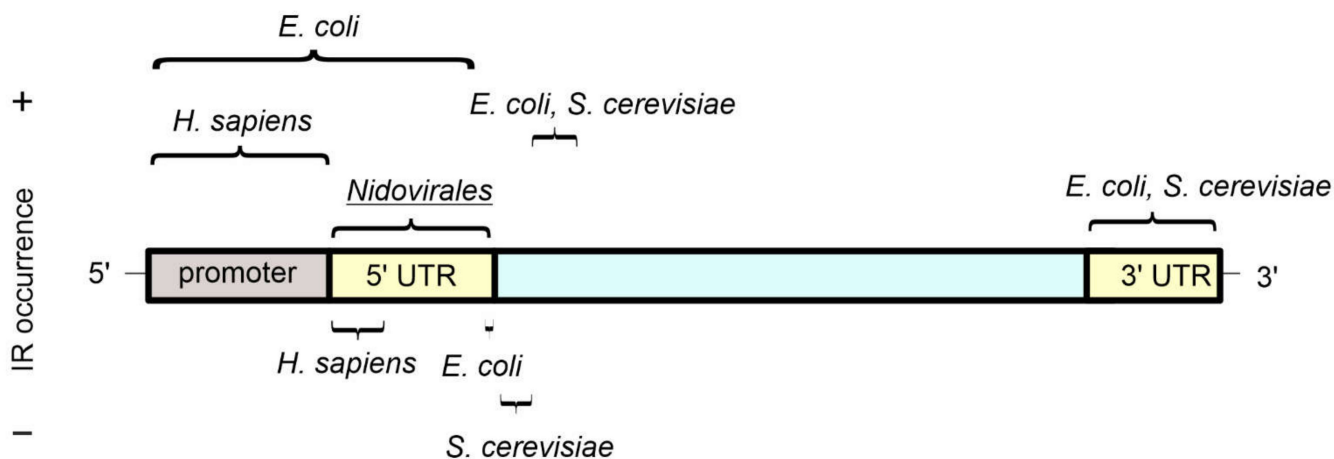


Figure 3. The occurrence of inverted repeat sequences in gene features as determined by bioinformatic analyses. An idealised gene and its regulatory sequences are shown, with UTR referring to “untranslated regions”. A relative abundance (+) or depletion (−) of inverted repeats in the indicated genomes is highlighted above and below the idealised gene, respectively. For *E. coli* and *S. cerevisiae*, inverted repeats with a stem length from 5 bp and a spacer length up to 8 bp were considered [50,51], while for *H. sapiens* and viruses from the *Nidovirales* order, inverted repeats with the stem length from 6–30 bp and spacer length up to 10 bp were taken into account [18,42].

3.2. Prokaryotes

Early evidence for the presence of inverted repeats and cruciforms in genomes was obtained from studies across a range of bacteria, with a particular focus on *E. coli* [6]. Because bacterial DNA is often circular, it easily results in a negative supercoiled conformation [52], which can be an important factor in the formation of cruciforms. In the *E. coli* genome,

short inverted repeats with arm lengths from 5 bp up to 20 bp are abundant in both coding and non-coding regions [19]. On average, there are nine inverted repeats per non-coding region, although a small proportion of regions contain the majority of the inverted repeats. The average arm length of the inverted repeats is approximately 6 bp, suggesting the sequences can form stable cruciforms. When comparing the genome with other proteobacteria, a significant number of identical inverted repeats are observed, providing evidence for evolutionary conservation [19]. Another study of genome sequences [53] performed similar analyses on 37 genomes of various prokaryotes, namely archaea, chlamydiales, firmicutes, proteobacteria, and others. For all bacteria, inverted repeats were found more frequently in non-coding regions. In almost all bacterial species examined, inverted repeats were found in genomes at a significantly higher frequency than the randomly generated sequences. Notably, only in two species, *Deinococcus radiodurans* and *Synechocystis* sp., were the observed number of inverted repeats statistically significantly lower than predicted by Markovian models of DNA sequences, although the reasons for the differences in these genomes are unclear. In archaea, the frequencies were higher than expected for five of eight species that were studied, but even in the five species that were higher, the difference was relatively small compared to that seen for bacteria. Mapping of the occurrence of inverted repeats in the *E. coli* genome [50] found that sequences with the potential to form cruciforms are enriched near stop codons and are part of terminators—and thus probably serve in the Rho-independent termination of transcription (Figure 3). Inverted repeats are also enriched within promoters, 5'-untranslated regions (UTRs), and in regions ~25–45 bp encompassing the start codon. It was also found that the small region ~5bp before the start codon has a statistically significant depletion of inverted repeats compared to 50 randomized genomes. Explanations for this observation could be that such a depletion prevents the formation of hairpin structures on the corresponding mRNA strands and also prevents disruption of the Shine Dalgarno sequence, both of which could negatively impact the initiation of translation.

For organisms that had complete genome sequences in 2020, about 36% of all bacteria and 75% of archaea have a prokaryotic immune system known as CRISPR/Cas [54]. CRISPR is an acronym for segments of clustered regularly interspaced short palindromic repeats, while Cas is the name of a group of proteins that associate with these regions. As the name implies, this system consists of sequences of inverted repeats, which are preceded by a leader sequence that is rich in adenine and thymine, and new spacers are integrated in its vicinity [55]. The nucleases Cas1 and Cas2 are the only Cas proteins that occur in all CRISPR/Cas systems, and both nucleases require a negatively supercoiled conformation to integrate new intervening sequences [56,57]. In vitro, the Cas1-Cas2 complex is able to integrate the new intervening sequence outside the CRISPR locus; however, the integration is non-random. In studying the specificity of integration of new intervening sequences, it was found that in the absence of the CRISPR locus, integration occurred preferentially in the vicinity of inverted repeats capable of forming cruciforms [56]. The CRISPR/Cas methodology is gaining widespread use across all organisms, but the potential impact of cruciforms on its implementation requires further analyses.

3.3. Eukaryotes

In eukaryotes, inverted repeats occur frequently in nuclear DNA and also in mitochondrial and plastid DNA, usually in even higher numbers than in nuclear DNA [20,21,58]. For example, in *Saccharomyces cerevisiae*, inverted repeats in mitochondrial DNA are 45-times more frequent than in its chromosomal DNA [17]. Correspondingly, inverted repeats have been demonstrated to impact evolution in mitochondria and in other genome contexts [59,60]. An overlay with annotated features revealed a statistically significant deficiency of inverted repeats in regions 20 bp downstream of the start codon [51]. In a similar way to examples already discussed for *E. coli* [50], inverted repeats in *S. cerevisiae* are enriched in the region ~30–60 bp downstream of the start codon and in close vicinity of positions corresponding to the ends of the mRNA (Figure 3) [51]. Whereas inverted repeats

in *E. coli* are parts of intrinsic terminators and are GC-rich, inverted repeats in *S. cerevisiae* are parts of the polyA signal and are AT-rich. Therefore, inverted repeats in both organisms appear to play roles in transcription termination, although the sequences of the repeats are not preserved [50,51].

The effort to complete the sequence of the human genome is now successfully finished [61], with two chromosomes (8 and X) fully assembled already in 2021 [62,63]. Regions in chromosomes 8 and X that were uncharacterized in the current reference human genome assembly GRCh38 are now resolved and reveal a previous strong underestimation of the frequency of repeat tracts [64,65]. The difference of inverted repeat frequency between the two assemblies of chromosome 8 increases with the length of the inverted repeat, with up to twice as many for inverted repeats with an arm length of 30 bp [64]. When examining inverted repeats in promoters of the human genome [18], it was found that their frequency depends on the length of the repeat and its distance from the transcription start site. Shorter inverted repeats (6–11 bases for the size of the stem) are found primarily near the transcription start site, while longer repeats (14 bases and above for the size of the stem) are more frequent in regions that are at least 500 bp upstream from the transcription start site. In general, inverted repeats in the human genome are abundant upstream from the transcription start site, while downstream (in the direction of transcription), their presence is rarer (Figure 3). Some evidence suggests DNA is negatively supercoiled upstream of RNA polymerase [66], which will facilitate DNA strand separation and increase the likelihood that inverted repeats could form cruciforms [67]. The increased incidence of inverted repeats upstream of the transcription site would be consistent with these repeat sequences being involved in organizing and controlling promoter activities whether or not they form cruciforms [18,68]. It is also likely that the inverted repeats or potential cruciforms may impact differently on different transcription factors, as evidenced by promoters of genes involved in inflammatory, tumour, and developmental processes containing relatively high levels of inverted repeats, whereas promoters of metabolic-related genes contain lower levels of inverted repeats [18].

4. A Range of Proteins Interact with Cruciforms

Inverted repeats and cruciforms are targets for binding by many architectural and regulatory proteins. While many proteins have only weak sequence specificity, they are able to bind strongly to non-B-DNA structures, such as cruciforms [15]. Additionally, some proteins induce or stabilize cruciforms after binding to the nucleic acid. Cruciform binding proteins have been shown to have roles in chromatin remodelling, replication, and transcription regulation. Table 1 highlights the names and sources of proteins confirmed to interact with cruciforms, and details about the impact of some of these interactions have been discussed previously [15]. More recent findings in relation to the involvement of these interactions across the full range of cellular processes are described below.

Table 1. Proteins involved in interactions with cruciform structures. TF = transcription factor, chromatin AP = chromatin-associated protein. Adapted from [15]. * If no reference is listed for an entry, see [15] for further details.

Protein	Source	Function	Reference *
14-3-3	Eukaryotes	Replication, DNA repair, TF	[69]
A22	Coccinia virus	Junction-resolving enzyme	
AF10	<i>H. sapiens</i>	TF	
Bmh1, homolog of 14-3-3	<i>S. cerevisiae</i>	Replication, DNA repair, TF	
BRCA1	Mammals	Chromatin AP, DNA repair, TF	[70]
Cas1, Cas2	Archaea, Bacteria	Endonuclease, defence response to virus	[56,57]
Cce1	Yeast	Junction-resolving enzyme	[71]
Crp-1	<i>S. cerevisiae</i>	DNA repair	[72]
DEK	Mammals	Chromatin AP, replication, DNA repair	[73,74]
DNA-PK	Eukaryotes	DNA repair	

Table 1. Cont.

Protein	Source	Function	Reference *
Dps	<i>E. coli</i>	DNA repair, stress response	[75–77]
Endonuclease I	Phage T7	Junction-resolving enzyme	[78]
Endonuclease VII	Phage T4	Junction-resolving enzyme	
Estrogen receptor	Mammals	TF	
GEN1	Vertebrates	Junction-resolving enzyme	[79]
GF14, homolog of 14-3-3	Plants	Replication, stress response	
Helicases	all	Replication	[80,81]
Hjc, Hje	Archaea	Junction-resolving enzymes	
HMG protein family	all	Chromatin AP, DNA repair, TF	
Hop1	<i>S. cerevisiae</i>	DNA Repair	
HU	<i>E. coli</i>	Replication	[82]
IFI16	<i>H. sapiens</i>	Viral DNA recognition	[83,84]
Integrases	all	Junction-resolving enzyme	
MLH1-MLH3	Vertebrates	Junction-resolving enzyme	[85]
MLL (leukaemia)	<i>H. sapiens</i>	Replication	
MSH2	Mammals	Junction-resolving enzyme	[86]
Mus81-Eme1	Eukaryotes	Junction-resolving enzyme	
Mus81-Mms4	<i>S. cerevisiae</i>	Junction-resolving enzyme	[72,87]
MutH	Eukaryotes	Junction-resolving enzyme	
p53	<i>H. sapiens</i> and others	DNA repair, TF	[88]
p73	<i>H. sapiens</i> and others	DNA repair, TF	[89]
PARP-1	<i>H. sapiens</i> and others	DNA repair, TF	[90]
Rad51	Eukaryotes	Chromatin AP	[91]
Rad52-Rad59	Eukaryotes	Chromatin AP	[91]
Rad54	Eukaryotes	Chromatin AP	[91]
RecU	G+ bacteria	Junction-resolving enzyme	
RepC	Bacteria	Replication	[92]
Rif1	Mammals	DNA repair, TF	[93,94]
Rmi-1	Yeast	DNA repair, TF	
RusA	<i>E. coli</i>	Junction-resolving enzyme	
RuvC	<i>E. coli</i>	Junction-resolving enzyme	
S16	<i>E. coli</i>	Replication	
SbcCD	<i>E. coli</i>	Junction-resolving enzyme	[95]
Smc	<i>S. cerevisiae</i>	DNA repair, TF	
Topoisomerase I	Eukaryotes	Chromatin AP	
Topoisomerase II	Eukaryotes	Chromatin AP	[96]
TRF2	<i>H. sapiens</i>	Junction-resolving enzyme	
Vlf-1	Baculoviruses	Replication	
WRN(Werner syndrome)	<i>H. sapiens</i>	Replication	
XPF, XPG protein families	Eukaryotes	Junction-resolving enzyme	[97]
Ydc2	<i>S. pombe</i>	Junction-resolving enzyme	
Yen1, homolog of GEN1	<i>S. cerevisiae</i>	Junction-resolving enzyme	[98]

Cruciform formation is enabled by DNA negative supercoiling, which is unevenly spread through genomes and is tightly regulated, mainly by topoisomerases (TOPs) [15]. In eukaryotes, TOP1 relaxes DNA supercoiling generated by transcription, replication, and chromatin remodelling through the introduction of a single-strand break, and it binds to Holliday junctions, whereas TOP2 changes the DNA topology and is capable of generating transient DNA double-strand breaks [99]. TOP2 has been shown to recognize and cleave cruciform structures [15]. TOP2 and a member of the HMG family, chromatin-stabilizing protein Hmo1, preserve negative supercoiling at gene boundaries and are suggested to instigate the formation of cruciforms, thus directing TOP1 and RNA polymerase II to coding regions [96].

Inverted repeats located in the promoter regions of genes are preferentially bound by many transcription factors (Table 1), such as PARP-1, BRCA1, ER, and p53 [15,70,90]. The tumour suppressor protein p53 is critical for protection against many human cancers.

Most tumorigenic p53 mutations occur in its central domain, which binds to specific DNA sequences, referred to as response elements. Such response elements with a propensity to form cruciforms are favoured for binding by p53 both in vitro and in vivo [14,100]. The protein p73 is a member of the p53 family and has essential functions in several signaling pathways involved in development, differentiation, DNA damage responses, and cancer. Like its p53 homolog, p73 shows a preference for binding to its target sequence in cruciform structures [89]. Yeast-based assays revealed that p73-mediated transactivation correlated with the relative propensity of a response element to form a cruciform [89].

Another protein showing a preference for binding to DNA cruciforms is interferon-inducible protein 16 (IFI16), a sensor of foreign DNA in human cells. Upon DNA recognition, the protein oligomerizes, forms a filament, and triggers an innate immune response [101]. Besides its role in the immune response, IFI16 represses the transcription of viral genes [102]. IFI16 showed a preference for binding to negatively supercoiled plasmid over linear DNA in vitro, stabilizing local DNA structures such as cruciforms and quadruplexes [83]. Importantly, the binding pattern varies dependent on secondary structures in the DNA: with linear DNA, the protein interacts cooperatively, leading to non-specific filamentous aggregates of a higher molecular weight being formed, but in the presence of cruciforms, the protein binds to DNA selectively, forming more compact globular complexes [83,84]. The functional role of the different binding patterns remains unclear, but provides a possible explanation for the distinct roles of IFI16 in antiviral defence.

Cruciforms have also been demonstrated to influence various aspects of DNA replication. A range of studies confirmed cruciform formation in the origins of replication in bacteria, yeast, and mammalian cells [15,103]. Furthermore, several proteins involved in replication bind to cruciform structures, such as S16, MLL, WRN, and 14-3-3 (Table 1). Replication initiator protein C (RepC), which is encoded by the pT181 plasmid of *Staphylococcus aureus*, binds to a specific DNA sequence, which is able to form a cruciform and creates a nick that allows replication to begin [104]. It is proposed that cruciforms are formed passively due to the natural supercoiling of DNA, but their formation is necessary for RepC cleavage of DNA [92]. Rap1-interacting factor 1 (Rif1) is a mammalian protein involved in regulating the timing of DNA replication, mediating the repair of double-stranded DNA breaks, and replication fork restart [93]. The C-terminal region CII of RIF1 is critical for replication fork protection, and recent structural analyses identified that it preferentially binds cruciform structures [93,94]. Rif1 accumulates on stalled replication forks and possibly protects reversed forks, which could involve cruciform structures in vivo.

Cruciforms also influence other aspects of replication. Cruciforms formed ahead of a replication fork could stop their movement, which would temporarily stop replication. Such problems can be resolved by the formation of reversed replication forks at the four-way junctions, followed by homologous recombination and branch migration in order to restart replication [105]. Since cruciforms share structural similarity with Holliday junctions, cruciform-binding proteins are likely to be involved in these (or related) processes. For example, AT-rich cruciform cleavage is mediated by the Holliday junction resolvase GEN-1 in human cells [79,106], with GEN1 splitting the cruciform diagonally, creating two hairpins healed by DNA ligases [79]. The tips of these hairpins are then cleaved by Artemis proteins and joined by non-homologous end joining. The resulting heteroduplexes are repaired by proteins associated with mismatch repair (MMR), for which the template would normally be selected according to the strand where the nick is not ligated. Since, however, both strands are fully ligated, the template is chosen randomly and may result in translocation between two palindromic AT-rich repeats at different chromosomal locations that do not share a complete sequence homology. The involvement of other resolvases in this type of process, such as Mus81 in human cells, was rejected [79]. However, in *S. cerevisiae*, Mus81-Mms4 was able to process recombination intermediates that arose during the repair of stalled replication forks and double-stranded breaks after being stimulated by Crp1, a protein that specifically binds to DNA four-way junctions [72,107].

Notably, long inverted repeats with an arm length of more than 150–200 nucleotides and with a spacer between the repeats being shorter than 50–60 nucleotides are almost impossible to clone into *E. coli*, mainly due to the action of SbcCD endonuclease/exonuclease, which can cleave hairpin structures, leading to DNA double-strand breaks [95,108]. It was confirmed that such long inverted repeats are converted to cruciform DNA before they encounter the replication fork, creating SbcCD-sensitive hairpin structures on both leading and lagging strands that transiently impede replication fork movement [109].

Another example of a protein able to bind to cruciforms is DNA-binding protein from starved cells (Dps), which is produced in stationary-phase *E. coli* cells on a large scale, reaching 85,000–180,000 molecules per cell. The main role of Dps is to protect cells from oxidative stress, UV- and γ -radiation, and metal ion toxicity, which it does via its ferroxidase activity [75]. Dps also regulates transcription by competing for binding sites with other transcription factors [76]. Dps protein binding to DNA does not depend on sequence, but a non-random distribution of Dps binding sites was observed with significant correlation with inverted repeats, suggesting the protein may interact with specific structures in DNA [76,77].

5. Inverted Repeats and Cruciforms as Potential Therapeutics in Human Disease

Evidence presented so far clearly demonstrates that cruciforms can form within DNA molecules in cells and that proteins bind to them, but the physiological significance of these observations remains unclear, particularly for human cells. However, a recent analysis of 1000 human genomes estimated that the probability of occurrence of pathology-associated single-nucleotide polymorphism variants is 14-times higher in inverted repeats than in other genome sites [110], and their role has been shown in germline mutagenesis with implications for evolution and genetic diseases [111]. Single-nucleotide polymorphism variants in inverted repeats have been linked with many human neuronal disorders, mental retardation, and various cancers. Moreover, when amplified genomic regions are determined for various cancer types [112], short palindromes are observed to facilitate these processes and lead to cancer progression [113]. Due to the presence of inverted repeats in multiple parts of genomes that are associated with regulatory functions, cruciforms are likely to be involved in several basic biological processes with physiological and pathological importance (Figure 4).

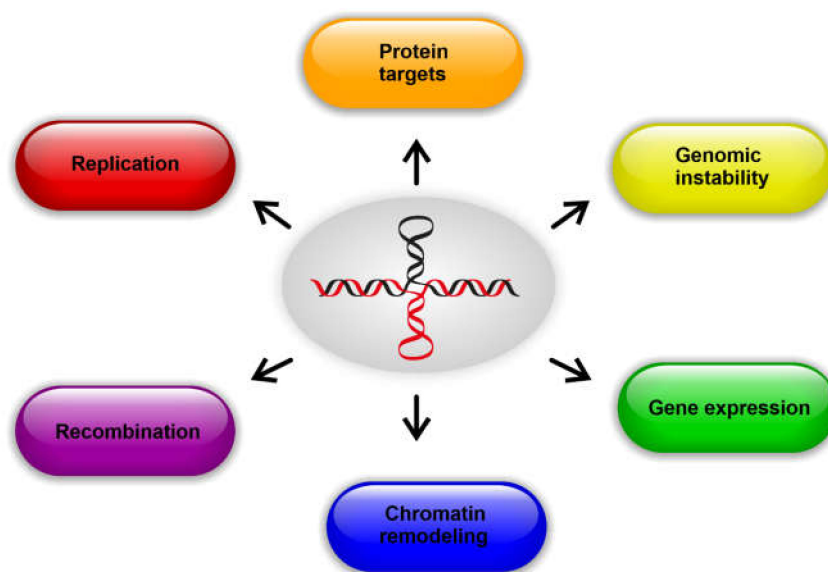


Figure 4. Cellular processes influenced by cruciform structures.

A range of local DNA structures are suggested as good therapeutic targets for human disease [9,114]. Considering that cruciforms formed by inverted repeats are hotspots of DNA breakpoints and for mutations with various pathologies [27,48], the detailed knowl-

edge presented within this review provides an important background for their use as therapeutic targets. Incomplete assemblies of genomes present significant problems in that sequences with good potential to form local DNA structures are often not characterized properly, and until recently, many repeat tracts have not been identified because sequencing technologies have not been able to cope with them [64]. Fortunately, contemporary sequencing technologies allow the complete assembly of even very complex genomes, including the human genome [63]. As described above, recent data of complete human chromosomes identified inverted repeats in the human genome that had previously not been seen [64]. The improved understanding of the widespread nature of these regulatory sequences will make it possible to judge more accurately whether their targeting is feasible for specific human diseases.

The range of structures that can be adopted within DNA have important impacts on genome integrity and genome plasticity. Thus, it is not surprising that cruciforms (and four-way junctions) play critical roles in the maintenance of genomic stability, with a concomitant impact on essential cellular processes [15,115]. For example, this is observed directly through their identification as hotspots of genomic rearrangements [115]. Molecular mechanisms have been inferred for how these types of structures mediate such rearrangements in the human genome [116], such as by Holliday junction resolvases mediating chromosomal translocations, as discussed above. Inverted repeats are frequently found at fragile sites in the genome that are prone to chromosome breakage, as shown for the fragile site FRA16D, where a variable-length AT repeat forms a cruciform that stalls replication [117]. The relative position and size of inverted repeats is also important in relation to their effects on genome stability. These parameters impact translocation frequency, with an inverted repeat arm size of up to 100 bp correlating with translocation breakpoints in human cancer genomes [97]. The involvement of structure-specific nucleases on the fragility of inverted repeats also depends on the distance between them and their transcriptional status [87]. The association of several human diseases with mutations of DNA helicases has also suggested possible roles for cruciforms in the diseases [118]. Although cruciforms may be important for basic biological processes, if they are not resolved by helicases, their presence could lead to transcription stop or delay and to chromosome breakage during replication. Dysfunction of these helicases can lead to various diseases, for example Werner's syndrome, which is associated with mutations in the WRN helicase [119]. Inverted repeats also play a key role in the transposition and reorganization of transposable elements as demonstrated in several disease models, for example in Williams–Beuren syndrome, where insertions and deletions are associated with genomic regions that have an abundant number of inverted repeats [120].

Cruciforms are already used for various applications in medicine. For example, a cruciform DNA nanostructure is used for targeted delivery of doxorubicin to cancer cells [121] and was used to treat colon cancer [122]. It has also been demonstrated that cruciforms in gene promoters impact transcription upon oxidative modification of 2'-Deoxyguanosine [123]. The association of cruciforms with the regulation of transcription [90], as discussed above, opens other therapeutic windows where the specific levels of gene expression are influenced by their presence and stability in promoter regions. An important tool allowing such approaches is the monoclonal antibody with specificity to the cruciform structure, although up to now, this has only been used for research purposes [28,69,124]. Currently, there are no small molecules that specifically recognize cruciforms, but it is likely that compounds will soon be designed that impact cruciform–protein interactions.

6. Conclusions

DNA molecules that contain inverted repeat sequences are able to adopt fully base-paired “linear” conformations and cruciforms that contain several unpaired regions. The structures of cruciforms (and four-way junctions) have been best characterized *in vitro*, including in complexes with proteins from prokaryotes and eukaryotes that bind to hairpins

and four-way junctions. The structure of the cruciform influences the thermodynamic stability of the DNA, and paired regions of at least 6 bp are usually required to offset the energetically unfavourable folding of the junction and loop regions. In recent years, significant advances have been made in identifying high-resolution analyses of unusual DNA structures, either as the nucleic acid alone or in combination with proteins. A range of structural studies has demonstrated that cruciforms and four-way junctions adopt different structures depending on the ionic environment and other factors, including whether or not proteins are interacting with them. High-resolution structures that are currently available for four-way DNA junctions are usually for sequences that are not inverted repeats, but it is expected that structures formed by inverted repeats will have similar flexibility, although this still has to be verified by high-resolution structures. It will be useful to confirm at high resolution whether proteins bind to the junction, stem, or loops, or whether this is protein-dependent.

Detailed studies of many organisms have identified that inverted repeats are widespread in natural genomes. Indeed, in most cases, they are found at higher levels than expected if these were present at just random frequencies. This suggests that these types of sequences and/or their structures have functions in cells. In most eukaryotes, inverted repeats occur in higher amounts near promoters and transcriptional terminators, whereas in prokaryotes, they occur more frequently close to terminators. Both observations suggest these sequences and/or their cruciform structures have roles in regulating transcription. A similar increase in the amount of inverted repeats occurs near the origins of replication in eukaryotes, suggesting that the proteins involved in the initiation of replication may bind to these sequences and/or the structures within them.

The presence of inverted repeats can have negative effects on genome stability, and they have been shown to promote mutations and are, thus, an important driver of evolution. When examined in relation to human diseases, such as a range of cancers, genetic rearrangements are often abundant and complex, meaning it can be difficult to unravel the events that start and then lead to a certain genotype. Clearly, amplifications of inverted repeats have important impacts on the mechanisms involved in carcinogenesis, but their exact roles in diseases remain unclear; those that exist in the human genome could have a much greater role in initiating recombination events than is currently appreciated.

Although inverted repeats have been the subject of many studies over the last 50 years, their distribution has recently become an increased focus of research due to developments in sequencing and computer software. It is now clear that inverted repeats are conserved and not randomly distributed in genomes, suggesting that they play important roles in nucleic acid metabolism. In the future, advances with *in vitro* and *in vivo* methods will allow experimental examination of the predictions from bioinformatics analyses, facilitating thorough investigations into the effects of cruciforms on cellular processes, providing a deeper understanding of the resulting effects on human disease.

Author Contributions: Conceptualization, R.P.B. and V.B.; formal data and literature analysis, N.B. and V.B.; writing—original draft preparation, review and editing, R.P.B., N.B. and V.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by The Czech Science Foundation (No. 22-21903S) to V.B.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

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

References

1. Sato, M.P.; Ogura, Y.; Nakamura, K.; Nishida, R.; Gotoh, Y.; Hayashi, M.; Hisatsune, J.; Sugai, M.; Takehiko, I.; Hayashi, T. Comparison of the Sequencing Bias of Currently Available Library Preparation Kits for Illumina Sequencing of Bacterial Genomes and Metagenomes. *DNA Res.* **2019**, *26*, 391–398. [[CrossRef](#)] [[PubMed](#)]
2. Oprzeska-Zingrebe, E.A.; Meyer, S.; Roloff, A.; Kunte, H.-J.; Smiatek, J. Influence of Compatible Solute Ectoine on Distinct DNA Structures: Thermodynamic Insights into Molecular Binding Mechanisms and Destabilization Effects. *Phys. Chem. Chem. Phys.* **2018**, *20*, 25861–25874. [[CrossRef](#)] [[PubMed](#)]
3. Brazda, V.; Fojta, M.; Bowater, R.P. Structures and Stability of Simple DNA Repeats from Bacteria. *Biochem. J.* **2020**, *477*, 325–339. [[CrossRef](#)] [[PubMed](#)]
4. Summers, P.A.; Lewis, B.W.; Gonzalez-Garcia, J.; Porreca, R.M.; Lim, A.H.M.; Cadinu, P.; Martin-Pintado, N.; Mann, D.J.; Edel, J.B.; Vannier, J.B.; et al. Visualising G-Quadruplex DNA Dynamics in Live Cells by Fluorescence Lifetime Imaging Microscopy. *Nat. Commun.* **2021**, *12*, 162. [[CrossRef](#)]
5. Di Antonio, M.; Ponjavic, A.; Radzevičius, A.; Ranasinghe, R.T.; Catalano, M.; Zhang, X.; Shen, J.; Needham, L.-M.; Lee, S.F.; Klenerman, D.; et al. Single-Molecule Visualization of DNA G-Quadruplex Formation in Live Cells. *Nat. Chem.* **2020**, *12*, 832–837. [[CrossRef](#)]
6. Poggi, L.; Richard, G.-F. Alternative DNA Structures In Vivo: Molecular Evidence and Remaining Questions. *Microbiol. Mol. Biol. Rev.* **2021**, *85*, e00110-20. [[CrossRef](#)]
7. Brown, J.A. Unraveling the Structure and Biological Functions of RNA Triple Helices. *Wiley Interdiscip. Rev. RNA* **2020**, *11*, e1598. [[CrossRef](#)]
8. Neil, A.J.; Liang, M.U.; Khristich, A.N.; Shah, K.A.; Mirkin, S.M. RNA–DNA Hybrids Promote the Expansion of Friedreich’s Ataxia (GAA)n Repeats via Break-Induced Replication. *Nucleic Acids Res.* **2018**, *46*, 3487–3497. [[CrossRef](#)]
9. Kosiol, N.; Juranek, S.; Brossart, P.; Heine, A.; Paeschke, K. G-Quadruplexes: A Promising Target for Cancer Therapy. *Mol. Cancer* **2021**, *20*, 40. [[CrossRef](#)]
10. Martella, M.; Pichiorri, F.; Chikhale, R.V.; Abdelhamid, M.A.S.; Waller, Z.A.E.; Smith, S.S. I-Motif Formation and Spontaneous Deletions in Human Cells. *Nucleic Acids Res.* **2022**, *50*, gkac158. [[CrossRef](#)]
11. Niehrs, C.; Luke, B. Regulatory R-Loops as Facilitators of Gene Expression and Genome Stability. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 167–178. [[CrossRef](#)]
12. Tye, S.; Ronson, G.E.; Morris, J.R. A Fork in the Road: Where Homologous Recombination and Stalled Replication Fork Protection Part Ways. *Semin. Cell Dev. Biol.* **2021**, *113*, 14–26. [[CrossRef](#)] [[PubMed](#)]
13. Palecek, E. Local Supercoil-Stabilized DNA Structures. *Crit. Rev. Biochem. Mol. Biol.* **1991**, *26*, 151–226. [[CrossRef](#)]
14. Brázda, V.; Fojta, M. The Rich World of P53 DNA Binding Targets: The Role of DNA Structure. *Int. J. Mol. Sci.* **2019**, *20*, 5605. [[CrossRef](#)] [[PubMed](#)]
15. Brázda, V.; Laister, R.C.; Jagelská, E.B.; Arrowsmith, C. Cruciform Structures Are a Common DNA Feature Important for Regulating Biological Processes. *BMC Mol. Biol.* **2011**, *12*, 33. [[CrossRef](#)] [[PubMed](#)]
16. Benham, C.J.; Savitt, A.G.; Bauer, W.R. Extrusion of an Imperfect Palindrome to a Cruciform in Superhelical DNA: Complete Determination of Energetics Using a Statistical Mechanical Model. *J. Mol. Biol.* **2002**, *316*, 563–581. [[CrossRef](#)] [[PubMed](#)]
17. Čutová, M.; Manta, J.; Porubiaková, O.; Kaura, P.; Šťastný, J.; Jagelská, E.B.; Goswami, P.; Bartas, M.; Brázda, V. Divergent Distributions of Inverted Repeats and G-Quadruplex Forming Sequences in *Saccharomyces Cerevisiae*. *Genomics* **2020**, *112*, 1897–1901. [[CrossRef](#)]
18. Brázda, V.; Bartas, M.; Lýsek, J.; Coufal, J.; Fojta, M. Global Analysis of Inverted Repeat Sequences in Human Gene Promoters Reveals Their Non-Random Distribution and Association with Specific Biological Pathways. *Genomics* **2020**, *112*, 2772–2777. [[CrossRef](#)]
19. Lavi, B.; Karin, E.L.; Pupko, T.; Hazkani-Covo, E. The Prevalence and Evolutionary Conservation of Inverted Repeats in Proteobacteria. *Genome Biol. Evol.* **2018**, *10*, 918–927. [[CrossRef](#)]
20. Brázda, V.; Lýsek, J.; Bartas, M.; Fojta, M. Complex Analyses of Short Inverted Repeats in All Sequenced Chloroplast DNAs. *BioMed Res. Int.* **2018**, *2018*, 1097018. [[CrossRef](#)]
21. Čechová, J.; Lýsek, J.; Bartas, M.; Brázda, V. Complex Analyses of Inverted Repeats in Mitochondrial Genomes Revealed Their Importance and Variability. *Bioinformatics* **2018**, *34*, 1081–1085. [[CrossRef](#)] [[PubMed](#)]
22. Gierer, A. Model for DNA and Protein Interactions and the Function of the Operator. *Nature* **1966**, *212*, 1480–1481. [[CrossRef](#)] [[PubMed](#)]
23. Murchie, A.I.; Bowater, R.; Aboul-ela, F.; Lilley, D.M. Helix Opening Transitions in Supercoiled DNA. *Biochim. Biophys. Acta BBA Gene Struct. Expr.* **1992**, *1131*, 1–15. [[CrossRef](#)]
24. Zhabinskaya, D.; Benham, C.J. Competitive Superhelical Transitions Involving Cruciform Extrusion. *Nucleic Acids Res.* **2013**, *41*, 9610–9621. [[CrossRef](#)] [[PubMed](#)]
25. Neelsen, K.J.; Chaudhuri, A.R.; Follonier, C.; Herrador, R.; Lopes, M. Visualization and Interpretation of Eukaryotic DNA Replication Intermediates In Vivo by Electron Microscopy. In *Functional Analysis of DNA and Chromatin*; Methods in Molecular Biology; Stockert, J.C., Espada, J., Blázquez-Castro, A., Eds.; Humana Press: Totowa, NJ, USA, 2014; pp. 177–208. ISBN 978-1-62703-706-8.

26. Torregrosa-Muñumer, R.; Goffart, S.; Haikonen, J.A.; Pohjoismäki, J.L.O. Low Doses of Ultraviolet Radiation and Oxidative Damage Induce Dramatic Accumulation of Mitochondrial DNA Replication Intermediates, Fork Regression, and Replication Initiation Shift. *Mol. Biol. Cell* **2015**, *26*, 4197–4208. [[CrossRef](#)] [[PubMed](#)]
27. Correll-Tash, S.; Lilley, B.; Iv, H.S.; Mlynarski, E.; Franconi, C.P.; McNamara, M.; Woodbury, C.; Easley, C.A.; Emanuel, B.S. Double Strand Breaks (DSBs) as Indicators of Genomic Instability in PATRR-Mediated Translocations. *Hum. Mol. Genet.* **2021**, *29*, 3872–3881. [[CrossRef](#)]
28. Rekvig, O.P. The Anti-DNA Antibodies: Their Specificities for Unique DNA Structures and Their Unresolved Clinical Impact—A System Criticism and a Hypothesis. *Front. Immunol.* **2022**, *12*, 808008. [[CrossRef](#)]
29. Brázda, V.; Kolomazník, J.; Lýsek, J.; Hároníková, L.; Coufal, J.; Št'astný, J. Palindrome Analyser—A New Web-Based Server for Predicting and Evaluating Inverted Repeats in Nucleotide Sequences. *Biochem. Biophys. Res. Commun.* **2016**, *478*, 1739–1745. [[CrossRef](#)]
30. Gibbs, D.R.; Dhakal, S. Homologous Recombination under the Single-Molecule Fluorescence Microscope. *Int. J. Mol. Sci.* **2019**, *20*, 6102. [[CrossRef](#)]
31. Stefanovsky, V.Y.; Moss, T. The Cruciform DNA Mobility Shift Assay: A Tool to Study Proteins That Recognize Bent DNA. *Methods Mol. Biol. Clifton NJ* **2015**, *1334*, 195–203. [[CrossRef](#)]
32. Ramreddy, T.; Sachidanandam, R.; Strick, T.R. Real-Time Detection of Cruciform Extrusion by Single-Molecule DNA Nanomanipulation. *Nucleic Acids Res.* **2011**, *39*, 4275–4283. [[CrossRef](#)] [[PubMed](#)]
33. Shaheen, C.; Hastie, C.; Metera, K.; Scott, S.; Zhang, Z.; Chen, S.; Gu, G.; Weber, L.; Munsky, B.; Kouzine, F.; et al. Non-Equilibrium Structural Dynamics of Supercoiled DNA Plasmids Exhibits Asymmetrical Relaxation. *Nucleic Acids Res.* **2022**, *50*, 2754–2764. [[CrossRef](#)] [[PubMed](#)]
34. Mandal, S.; Selvam, S.; Cui, Y.; Hoque, M.E.; Mao, H. Mechanical Cooperativity in DNA Cruciform Structures. *ChemPhysChem* **2018**, *19*, 2627–2634. [[CrossRef](#)] [[PubMed](#)]
35. Lilley, D.M.J. Holliday Junction-Resolving Enzymes—Structures and Mechanisms. *FEBS Lett.* **2017**, *591*, 1073–1082. [[CrossRef](#)]
36. Ho, P.S. Structure of the Holliday Junction: Applications beyond Recombination. *Biochem. Soc. Trans.* **2017**, *45*, 1149–1158. [[CrossRef](#)]
37. Eichman, B.F.; Vargason, J.M.; Mooers, B.H.M.; Ho, P.S. The Holliday Junction in an Inverted Repeat DNA Sequence: Sequence Effects on the Structure of Four-Way Junctions. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 3971–3976. [[CrossRef](#)]
38. Yadav, R.K.; Yadava, U. Molecular Dynamics Simulation of Hydrated d(CGCGTACCCG)₄ as a Four-Way DNA Holliday Junction and Comparison with the Crystallographic Structure. *Mol. Simul.* **2016**, *42*, 25–30. [[CrossRef](#)]
39. Kulkarni, D.S.; Owens, S.N.; Honda, M.; Ito, M.; Yang, Y.; Corrigan, M.W.; Chen, L.; Quan, A.L.; Hunter, N. PCNA Activates the MutLγ Endonuclease to Promote Meiotic Crossing Over. *Nature* **2020**, *586*, 623–627. [[CrossRef](#)]
40. Yan, J.; Hong, S.; Guan, Z.; He, W.; Zhang, D.; Yin, P. Structural Insights into Sequence-Dependent Holliday Junction Resolution by the Chloroplast Resolvase MOC1. *Nat. Commun.* **2020**, *11*, 1417. [[CrossRef](#)]
41. Wendorff, T.J.; Berger, J.M. Topoisomerase VI Senses and Exploits Both DNA Crossings and Bends to Facilitate Strand Passage. *eLife* **2018**, *7*, e31724. [[CrossRef](#)]
42. Bartas, M.; Brázda, V.; Bohálová, N.; Cantara, A.; Volná, A.; Stachurová, T.; Malachová, K.; Jagelská, E.B.; Porubiaková, O.; Červeň, J.; et al. In-Depth Bioinformatic Analyses of Nidovirales Including Human SARS-CoV-2, SARS-CoV, MERS-CoV Viruses Suggest Important Roles of Non-Canonical Nucleic Acid Structures in Their Lifecycles. *Front. Microbiol.* **2020**, *11*, 1583. [[CrossRef](#)] [[PubMed](#)]
43. Treangen, T.J.; Salzberg, S.L. Repetitive DNA and Next-Generation Sequencing: Computational Challenges and Solutions. *Nat. Rev. Genet.* **2012**, *13*, 36–46. [[CrossRef](#)]
44. Altemose, N.; Logsdon, G.A.; Bzikadze, A.V.; Sidhwani, P.; Langley, S.A.; Caldas, G.V.; Hoyt, S.J.; Uralsky, L.; Ryabov, F.D.; Shew, C.J.; et al. Complete Genomic and Epigenetic Maps of Human Centromeres. *Science* **2022**, *376*, eabl4178. [[CrossRef](#)] [[PubMed](#)]
45. Hoyt, S.J.; Storer, J.M.; Hartley, G.A.; Grady, P.G.S.; Gershman, A.; de Lima, L.G.; Limouse, C.; Halabian, R.; Wojenski, L.; Rodriguez, M.; et al. From Telomere to Telomere: The Transcriptional and Epigenetic State of Human Repeat Elements. *Science* **2022**, *376*, eabk3112. [[CrossRef](#)]
46. Spanò, M.; Lillo, F.; Miccichè, S.; Mantegna, R.N. Inverted Repeats in Viral Genomes. *Fluct. Noise Lett.* **2005**, *5*, L193–L200. [[CrossRef](#)]
47. Bartas, M.; Goswami, P.; Lexa, M.; Červeň, J.; Volná, A.; Fojta, M.; Brázda, V.; Pečinka, P. Letter to the Editor: Significant Mutation Enrichment in Inverted Repeat Sites of New SARS-CoV-2 Strains. *Brief. Bioinform.* **2021**, *22*, bbab129. [[CrossRef](#)]
48. Goswami, P.; Bartas, M.; Lexa, M.; Bohálová, N.; Volná, A.; Červeň, J.; Červeňová, V.; Pečinka, P.; Špunda, V.; Fojta, M.; et al. SARS-CoV-2 Hot-Spot Mutations Are Significantly Enriched within Inverted Repeats and CpG Island Loci. *Brief. Bioinform.* **2021**, *22*, 1338–1345. [[CrossRef](#)]
49. Berns, K.I. The Unusual Properties of the AAV Inverted Terminal Repeat. *Hum. Gene Ther.* **2020**, *31*, 518–523. [[CrossRef](#)]
50. Miura, O.; Ogake, T.; Ohyama, T. Requirement or Exclusion of Inverted Repeat Sequences with Cruciform-Forming Potential in Escherichia Coli Revealed by Genome-Wide Analyses. *Curr. Genet.* **2018**, *64*, 945–958. [[CrossRef](#)]
51. Miura, O.; Ogake, T.; Yoneyama, H.; Kikuchi, Y.; Ohyama, T. A Strong Structural Correlation between Short Inverted Repeat Sequences and the Polyadenylation Signal in Yeast and Nucleosome Exclusion by These Inverted Repeats. *Curr. Genet.* **2019**, *65*, 575–590. [[CrossRef](#)]

52. Lal, A.; Dhar, A.; Trostel, A.; Kouzine, F.; Seshasayee, A.S.N.; Adhya, S. Genome Scale Patterns of Supercoiling in a Bacterial Chromosome. *Nat. Commun.* **2016**, *7*, 11055. [[CrossRef](#)] [[PubMed](#)]
53. Lillo, F.; Basile, S.; Mantegna, R.N. Comparative Genomics Study of Inverted Repeats in Bacteria. *Bioinformatics* **2002**, *18*, 971–979. [[CrossRef](#)] [[PubMed](#)]
54. Pourcel, C.; Touchon, M.; Villeriot, N.; Vernadet, J.-P.; Couvin, D.; Toffano-Nioche, C.; Vergnaud, G. CRISPRCasdb a Successor of CRISPRdb Containing CRISPR Arrays and Cas Genes from Complete Genome Sequences, and Tools to Download and Query Lists of Repeats and Spacers. *Nucleic Acids Res.* **2020**, *48*, D535–D544. [[CrossRef](#)] [[PubMed](#)]
55. Makarova, K.S.; Grishin, N.V.; Shabalina, S.A.; Wolf, Y.I.; Koonin, E.V. A Putative RNA-Interference-Based Immune System in Prokaryotes: Computational Analysis of the Predicted Enzymatic Machinery, Functional Analogies with Eukaryotic RNAi, and Hypothetical Mechanisms of Action. *Biol. Direct* **2006**, *1*, 7. [[CrossRef](#)] [[PubMed](#)]
56. Nuñez, J.K.; Lee, A.S.Y.; Engelman, A.; Doudna, J.A. Integrase-Mediated Spacer Acquisition during CRISPR-Cas Adaptive Immunity. *Nature* **2015**, *519*, 193–198. [[CrossRef](#)] [[PubMed](#)]
57. Moch, C.; Fromant, M.; Blanquet, S.; Plateau, P. DNA Binding Specificities of Escherichia Coli Cas1-Cas2 Integrase Drive Its Recruitment at the CRISPR Locus. *Nucleic Acids Res.* **2017**, *45*, 2714–2723. [[CrossRef](#)]
58. Zhang, R.; Ge, F.; Li, H.; Chen, Y.; Zhao, Y.; Gao, Y.; Liu, Z.; Yang, L. PCIR: A Database of Plant Chloroplast Inverted Repeats. *Database J. Biol. Databases Curation* **2019**, 2019, baz127. [[CrossRef](#)]
59. Liu, X.; Wu, X.; Tan, H.; Xie, B.; Deng, Y. Large Inverted Repeats Identified by Intra-Specific Comparison of Mitochondrial Genomes Provide Insights into the Evolution of *Agrocybe Aegerita*. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 2424–2437. [[CrossRef](#)]
60. Damas, J.; Carneiro, J.; Gonçalves, J.; Stewart, J.B.; Samuels, D.C.; Amorim, A.; Pereira, F. Mitochondrial DNA Deletions Are Associated with Non-B DNA Conformations. *Nucleic Acids Res.* **2012**, *40*, 7606–7621. [[CrossRef](#)]
61. Nurk, S.; Koren, S.; Rhie, A.; Rautiainen, M.; Bizkadze, A.V.; Mikheenko, A.; Vollger, M.R.; Altemose, N.; Uralsky, L.; Gershman, A.; et al. The Complete Sequence of a Human Genome. *Science* **2022**, *376*, 44–53. [[CrossRef](#)]
62. Miga, K.H.; Koren, S.; Rhie, A.; Vollger, M.R.; Gershman, A.; Bizkadze, A.; Brooks, S.; Howe, E.; Porubsky, D.; Logsdon, G.A.; et al. Telomere-to-Telomere Assembly of a Complete Human X Chromosome. *Nature* **2020**, *585*, 79–84. [[CrossRef](#)] [[PubMed](#)]
63. Logsdon, G.A.; Vollger, M.R.; Hsieh, P.; Mao, Y.; Liskovych, M.A.; Koren, S.; Nurk, S.; Mercuri, L.; Dishuck, P.C.; Rhie, A.; et al. The Structure, Function and Evolution of a Complete Human Chromosome 8. *Nature* **2021**, *593*, 101–107. [[CrossRef](#)] [[PubMed](#)]
64. Brazda, V.; Bohalova, N.; Bowater, R.P. New Telomere to Telomere Assembly of Human Chromosome 8 Reveals a Previous Underestimation of G-Quadruplex Forming Sequences and Inverted Repeats. *Gene* **2021**, *810*, 146058. [[CrossRef](#)] [[PubMed](#)]
65. Bohálová, N.; Mergny, J.-L.; Brázda, V. Novel G-Quadruplex Prone Sequences Emerge in the Complete Assembly of the Human X Chromosome. *Biochimie* **2021**, *191*, 87–90. [[CrossRef](#)]
66. Forth, S.; Sheinin, M.Y.; Inman, J.; Wang, M.D. Torque Measurement at the Single Molecule Level. *Annu. Rev. Biophys.* **2013**, *42*, 583–604. [[CrossRef](#)]
67. Ma, J.; Wang, M.D. DNA Supercoiling during Transcription. *Biophys. Rev.* **2016**, *8*, 75–87. [[CrossRef](#)]
68. Yamamoto, Y.; Miura, O.; Ohyama, T. Cruciform Formable Sequences within Pou5f1 Enhancer Are Indispensable for Mouse ES Cell Integrity. *Int. J. Mol. Sci.* **2021**, *22*, 3399. [[CrossRef](#)]
69. Brázda, V.; Cechová, J.; Coufal, J.; Rumpel, S.; Jagelská, E.B. Superhelical DNA as a Preferential Binding Target of 14-3-3 γ Protein. *J. Biomol. Struct. Dyn.* **2012**, *30*, 371–378. [[CrossRef](#)]
70. Brázda, V.; Hároníková, L.; Liao, J.C.C.; Fridrichová, H.; Jagelská, E.B. Strong Preference of BRCA1 Protein to Topologically Constrained Non-B DNA Structures. *BMC Mol. Biol.* **2016**, *17*, 14. [[CrossRef](#)]
71. Samoilova, E.O.; Krashennnikov, I.A.; Levitskii, S.A. Interaction between *Saccharomyces Cerevisiae* Mitochondrial DNA-Binding Protein Abf2p and Cce1p Resolvase. *Biochemistry* **2016**, *81*, 1111–1117. [[CrossRef](#)]
72. Phung, H.T.T.; Tran, D.H.; Nguyen, T.X. The Cruciform DNA-Binding Protein Crp1 Stimulates the Endonuclease Activity of Mus81-Mms4 in *Saccharomyces Cerevisiae*. *FEBS Lett.* **2020**, *594*, 4320–4337. [[CrossRef](#)] [[PubMed](#)]
73. Deutzmann, A.; Ganz, M.; Schönenberger, F.; Vervoorts, J.; Kappes, F.; Ferrando-May, E. The Human Oncoprotein and Chromatin Architectural Factor DEK Counteracts DNA Replication Stress. *Oncogene* **2015**, *34*, 4270–4277. [[CrossRef](#)] [[PubMed](#)]
74. Martinez-Useros, J.; Rodriguez-Remirez, M.; Borrero-Palacios, A.; Moreno, I.; Cebrian, A.; del Pulgar, T.G.; del Puerto-Nevado, L.; Vega-Bravo, R.; Puime-Otin, A.; Perez, N.; et al. DEK Is a Potential Marker for Aggressive Phenotype and Irinotecan-Based Therapy Response in Metastatic Colorectal Cancer. *BMC Cancer* **2014**, *14*, 965. [[CrossRef](#)] [[PubMed](#)]
75. Calhoun, L.N.; Kwon, Y.M. Structure, Function and Regulation of the DNA-Binding Protein Dps and Its Role in Acid and Oxidative Stress Resistance in *Escherichia Coli*: A Review. *J. Appl. Microbiol.* **2011**, *110*, 375–386. [[CrossRef](#)]
76. Antipov, S.S.; Tutukina, M.N.; Preobrazhenskaya, E.V.; Kondrashov, F.A.; Patrushev, M.V.; Toshchakov, S.V.; Dominova, I.; Shvyreva, U.S.; Vrublevskaia, V.V.; Morenkov, O.S.; et al. The Nucleoid Protein Dps Binds Genomic DNA of *Escherichia Coli* in a Non-Random Manner. *PLoS ONE* **2017**, *12*, e0182800. [[CrossRef](#)]
77. Melekhov, V.V.; Shvyreva, U.S.; Timchenko, A.A.; Tutukina, M.N.; Preobrazhenskaya, E.V.; Burkova, D.V.; Artiukhov, V.G.; Ozoline, O.N.; Antipov, S.S. Modes of *Escherichia Coli* Dps Interaction with DNA as Revealed by Atomic Force Microscopy. *PLoS ONE* **2015**, *10*, e0126504. [[CrossRef](#)]
78. Freeman, A.D.J.; Déclais, A.-C.; Lilley, D.M.J. The Importance of the N-Terminus of T7 Endonuclease I in the Interaction with DNA Junctions. *J. Mol. Biol.* **2013**, *425*, 395–410. [[CrossRef](#)]

79. Inagaki, H.; Ohye, T.; Kogo, H.; Tsutsumi, M.; Kato, T.; Tong, M.; Emanuel, B.S.; Kurahashi, H. Two Sequential Cleavage Reactions on Cruciform DNA Structures Cause Palindrome-Mediated Chromosomal Translocations. *Nat. Commun.* **2013**, *4*, 1592. [[CrossRef](#)]
80. Li, D.; Lv, B.; Zhang, H.; Lee, J.Y.; Li, T. Disintegration of Cruciform and G-Quadruplex Structures during the Course of Helicase-Dependent Amplification (HDA). *Bioorg. Med. Chem. Lett.* **2015**, *25*, 1709–1714. [[CrossRef](#)]
81. Boyer, A.-S.; Grgurevic, S.; Cazaux, C.; Hoffmann, J.-S. The Human Specialized DNA Polymerases and Non-B DNA: Vital Relationships to Preserve Genome Integrity. *J. Mol. Biol.* **2013**, *425*, 4767–4781. [[CrossRef](#)]
82. Bettridge, K.; Verma, S.; Weng, X.; Adhya, S.; Xiao, J. Single-Molecule Tracking Reveals That the Nucleoid-Associated Protein HU Plays a Dual Role in Maintaining Proper Nucleoid Volume through Differential Interactions with Chromosomal DNA. *Mol. Microbiol.* **2021**, *115*, 12–27. [[CrossRef](#)] [[PubMed](#)]
83. Brázda, V.; Coufal, J.; Liao, J.; Arrowsmith, C. Preferential Binding of IFI16 Protein to Cruciform Structure and Superhelical DNA. *Biochem. Biophys. Res. Commun.* **2012**, *422*, 716–720. [[CrossRef](#)] [[PubMed](#)]
84. Hároníková, L.; Coufal, J.; Kejnovská, I.; Jagelská, E.B.; Fojta, M.; Dvořáková, P.; Muller, P.; Vojtesek, B.; Brázda, V. IFI16 Preferentially Binds to DNA with Quadruplex Structure and Enhances DNA Quadruplex Formation. *PLoS ONE* **2016**, *11*, e0157156. [[CrossRef](#)]
85. Cannavo, E.; Sanchez, A.; Anand, R.; Ranjha, L.; Hugener, J.; Adam, C.; Acharya, A.; Weyland, N.; Aran-Guiu, X.; Charbonnier, J.-B.; et al. Regulation of the MLH1–MLH3 Endonuclease in Meiosis. *Nature* **2020**, *586*, 618–622. [[CrossRef](#)] [[PubMed](#)]
86. Rogacheva, M.V.; Manhart, C.M.; Chen, C.; Guarne, A.; Surtees, J.; Alani, E. Mlh1-Mlh3, a Meiotic Crossover and DNA Mismatch Repair Factor, Is a Msh2-Msh3-Stimulated Endonuclease. *J. Biol. Chem.* **2014**, *289*, 5664–5673. [[CrossRef](#)]
87. Saada, A.A.; Costa, A.B.; Sheng, Z.; Guo, W.; Haber, J.E.; Lobachev, K.S. Structural Parameters of Palindromic Repeats Determine the Specificity of Nuclease Attack of Secondary Structures. *Nucleic Acids Res.* **2021**, *49*, 3932–3947. [[CrossRef](#)]
88. Brázda, V.; Čechová, J.; Battistin, M.; Coufal, J.; Jagelská, E.B.; Raimondi, I.; Inga, A. The Structure Formed by Inverted Repeats in P53 Response Elements Determines the Transactivation Activity of P53 Protein. *Biochem. Biophys. Res. Commun.* **2017**, *483*, 516–521. [[CrossRef](#)]
89. Čechová, J.; Coufal, J.; Jagelská, E.B.; Fojta, M.; Brázda, V. P73, like Its P53 Homolog, Shows Preference for Inverted Repeats Forming Cruciforms. *PLoS ONE* **2018**, *13*, e0195835. [[CrossRef](#)]
90. Feng, X.; Xie, F.-Y.; Ou, X.-H.; Ma, J.-Y. Cruciform DNA in Mouse Growing Oocytes: Its Dynamics and Its Relationship with DNA Transcription. *PLoS ONE* **2020**, *15*, e0240844. [[CrossRef](#)]
91. Marie, L.; Symington, L.S. Mechanism for Inverted-Repeat Recombination Induced by a Replication Fork Barrier. *Nat. Commun.* **2022**, *13*, 32. [[CrossRef](#)]
92. Pastrana, C.L.; Carrasco, C.; Akhtar, P.; Leuba, S.H.; Khan, S.A.; Moreno-Herrero, F. Force and Twist Dependence of RepC Nicking Activity on Torsionally-Constrained DNA Molecules. *Nucleic Acids Res.* **2016**, *44*, 8885–8896. [[CrossRef](#)] [[PubMed](#)]
93. Sukackaite, R.; Jensen, M.R.; Mas, P.J.; Blackledge, M.; Buonomo, S.B.; Hart, D.J. Structural and Biophysical Characterization of Murine Rif1 C Terminus Reveals High Specificity for DNA Cruciform Structures. *J. Biol. Chem.* **2014**, *289*, 13903–13911. [[CrossRef](#)] [[PubMed](#)]
94. Mukherjee, C.; Tripathi, V.; Manolika, E.M.; Heijink, A.M.; Ricci, G.; Merzouk, S.; de Boer, H.R.; Demmers, J.; van Vugt, M.A.T.M.; Chaudhuri, A.R. RIF1 Promotes Replication Fork Protection and Efficient Restart to Maintain Genome Stability. *Nat. Commun.* **2019**, *10*, 3287. [[CrossRef](#)] [[PubMed](#)]
95. Eykelenboom, J.K.; Blackwood, J.K.; Okely, E.; Leach, D.R.F. SbcCD Causes a Double-Strand Break at a DNA Palindrome in the Escherichia Coli Chromosome. *Mol. Cell* **2008**, *29*, 644–651. [[CrossRef](#)]
96. Achar, Y.J.; Adhil, M.; Choudhary, R.; Gilbert, N.; Foiani, M. Negative Supercoil at Gene Boundaries Modulates Gene Topology. *Nature* **2020**, *577*, 701–705. [[CrossRef](#)]
97. Lu, S.; Wang, G.; Bacolla, A.; Zhao, J.; Spitzer, S.; Vasquez, K.M. Short Inverted Repeats Are Hotspots for Genetic Instability: Relevance to Cancer Genomes. *Cell Rep.* **2015**, *10*, 1674–1680. [[CrossRef](#)]
98. Carreira, R.; Aguado, F.J.; Hurtado-Nieves, V.; Blanco, M.G. Canonical and Novel Non-Canonical Activities of the Holliday Junction Resolvase Yen1. *Nucleic Acids Res.* **2021**, *50*, 259–280. [[CrossRef](#)]
99. Vos, S.M.; Tretter, E.M.; Schmidt, B.H.; Berger, J.M. All Tangled up: How Cells Direct, Manage and Exploit Topoisomerase Function. *Nat. Rev. Mol. Cell Biol.* **2011**, *12*, 827–841. [[CrossRef](#)]
100. Coufal, J.; Jagelská, E.B.; Liao, J.C.C.; Brázda, V. Preferential Binding of P53 Tumor Suppressor to P21 Promoter Sites That Contain Inverted Repeats Capable of Forming Cruciform Structure. *Biochem. Biophys. Res. Commun.* **2013**, *441*, 83–88. [[CrossRef](#)]
101. Unterholzner, L.; Keating, S.E.; Baran, M.; Horan, K.A.; Jensen, S.B.; Sharma, S.; Sirois, C.M.; Jin, T.; Latz, E.; Xiao, T.S.; et al. IFI16 Is an Innate Immune Sensor for Intracellular DNA. *Nat. Immunol.* **2010**, *11*, 997–1004. [[CrossRef](#)]
102. Johnson, K.E.; Bottero, V.; Flaherty, S.; Dutta, S.; Singh, V.V.; Chandran, B. IFI16 Restricts HSV-1 Replication by Accumulating on the HSV-1 Genome, Repressing HSV-1 Gene Expression, and Directly or Indirectly Modulating Histone Modifications. *PLoS Pathog.* **2014**, *10*, e1004503. [[CrossRef](#)] [[PubMed](#)]
103. Toleikis, A.; Webb, M.R.; Molloy, J.E. OriD Structure Controls RepD Initiation during Rolling-Circle Replication. *Sci. Rep.* **2018**, *8*, 1206. [[CrossRef](#)] [[PubMed](#)]
104. Noirot, P.; Bargonetti, J.; Novick, R.P. Initiation of Rolling-Circle Replication in PT181 Plasmid: Initiator Protein Enhances Cruciform Extrusion at the Origin. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 8560–8564. [[CrossRef](#)]

105. Liao, H.; Ji, F.; Helleday, T.; Ying, S. Mechanisms for Stalled Replication Fork Stabilization: New Targets for Synthetic Lethality Strategies in Cancer Treatments. *EMBO Rep.* **2018**, *19*, e46263. [[CrossRef](#)]
106. Rass, U.; Compton, S.A.; Matos, J.; Singleton, M.R.; Ip, S.C.Y.; Blanco, M.G.; Griffith, J.D.; West, S.C. Mechanism of Holliday Junction Resolution by the Human GEN1 Protein. *Genes Dev.* **2010**, *24*, 1559–1569. [[CrossRef](#)] [[PubMed](#)]
107. Chen, S.; Geng, X.; Syeda, M.Z.; Huang, Z.; Zhang, C.; Ying, S. Human MUS81: A Fence-Sitter in Cancer. *Front. Cell Dev. Biol.* **2021**, *9*, 657305. [[CrossRef](#)] [[PubMed](#)]
108. Leach, D.R. Long DNA Palindromes, Cruciform Structures, Genetic Instability and Secondary Structure Repair. *Bioessays* **1994**, *16*, 893–900. [[CrossRef](#)] [[PubMed](#)]
109. Lai, P.J.; Lim, C.T.; Le, H.P.; Katayama, T.; Leach, D.R.F.; Furukohri, A.; Maki, H. Long Inverted Repeat Transiently Stalls DNA Replication by Forming Hairpin Structures on Both Leading and Lagging Strands. *Genes Cells* **2016**, *21*, 136–145. [[CrossRef](#)]
110. Ganapathiraju, M.K.; Subramanian, S.; Chaparala, S.; Karunakaran, K.B. A Reference Catalog of DNA Palindromes in the Human Genome and Their Variations in 1000 Genomes. *Hum. Genome Var.* **2020**, *7*, 40. [[CrossRef](#)]
111. Guiblet, W.M.; Cremona, M.A.; Harris, R.S.; Chen, D.; Eckert, K.A.; Chiaromonte, F.; Huang, Y.-F.; Makova, K.D. Non-B DNA: A Major Contributor to Small- and Large-Scale Variation in Nucleotide Substitution Frequencies across the Genome. *Nucleic Acids Res.* **2021**, *49*, 1497–1516. [[CrossRef](#)]
112. Tanaka, H.; Watanabe, T. Mechanisms Underlying Recurrent Genomic Amplification in Human Cancers. *Trends Cancer* **2020**, *6*, 462–477. [[CrossRef](#)] [[PubMed](#)]
113. Tanaka, H.; Tapscott, S.J.; Trask, B.J.; Yao, M.-C. Short Inverted Repeats Initiate Gene Amplification through the Formation of a Large DNA Palindrome in Mammalian Cells. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 8772–8777. [[CrossRef](#)] [[PubMed](#)]
114. Lopes-Nunes, J.; Oliveira, P.A.; Cruz, C. G-Quadruplex-Based Drug Delivery Systems for Cancer Therapy. *Pharmaceuticals* **2021**, *14*, 671. [[CrossRef](#)] [[PubMed](#)]
115. Miklenić, M.S.; Svetec, I.K. Palindromes in DNA—A Risk for Genome Stability and Implications in Cancer. *Int. J. Mol. Sci.* **2021**, *22*, 2840. [[CrossRef](#)]
116. Inagaki, H.; Kato, T.; Tsutsumi, M.; Ouchi, Y.; Ohye, T.; Kurahashi, H. Palindrome-Mediated Translocations in Humans: A New Mechanistic Model for Gross Chromosomal Rearrangements. *Front. Genet.* **2016**, *7*, 125. [[CrossRef](#)]
117. Kaushal, S.; Wollmuth, C.E.; Das, K.; Hile, S.E.; Regan, S.B.; Barnes, R.P.; Haouzi, A.; Lee, S.M.; House, N.C.M.; Guyumdzhyan, M.; et al. Sequence and Nuclease Requirements for Breakage and Healing of a Structure-Forming (AT)_n Sequence within Fragile Site FRA16D. *Cell Rep.* **2019**, *27*, 1151–1164.e5. [[CrossRef](#)]
118. Brosh, R.M., Jr.; Matson, S.W. History of DNA Helicases. *Genes* **2020**, *11*, 255. [[CrossRef](#)]
119. Datta, A.; Brosh, R.M., Jr. New Insights into DNA Helicases as Druggable Targets for Cancer Therapy. *Front. Mol. Biosci.* **2018**, *5*, 59. [[CrossRef](#)]
120. Savvateeva-Popova, E.V.; Zhuravlev, A.V.; Brázda, V.; Zakharov, G.A.; Kaminskaya, A.N.; Medvedeva, A.V.; Nikitina, E.A.; Tokmatcheva, E.V.; Dolgaya, J.F.; Kulikova, D.A.; et al. Drosophila Model for the Analysis of Genesis of LIM-Kinase 1-Dependent Williams-Beuren Syndrome Cognitive Phenotypes: INDELS, Transposable Elements of the Tc1/Mariner Superfamily and MicroRNAs. *Front. Genet.* **2017**, *8*, 123. [[CrossRef](#)]
121. Abnous, K.; Danesh, N.M.; Ramezani, M.; Charbgoor, F.; Bahreyni, A.; Taghdisi, S.M. Targeted Delivery of Doxorubicin to Cancer Cells by a Cruciform DNA Nanostructure Composed of AS1411 and FOXM1 Aptamers. *Expert Opin. Drug Deliv.* **2018**, *15*, 1045–1052. [[CrossRef](#)]
122. Yao, F.; An, Y.; Li, X.; Li, Z.; Duan, J.; Yang, X.-D. Targeted Therapy of Colon Cancer by Aptamer-Guided Holliday Junctions Loaded with Doxorubicin. *Int. J. Nanomed.* **2020**, *15*, 2119–2129. [[CrossRef](#)] [[PubMed](#)]
123. Fleming, A.M.; Zhu, J.; Jara-Espejo, M.; Burrows, C.J. Cruciform DNA Sequences in Gene Promoters Can Impact Transcription upon Oxidative Modification of 2'-Deoxyguanosine. *Biochemistry* **2020**, *59*, 2616–2626. [[CrossRef](#)] [[PubMed](#)]
124. Kurahashi, H.; Inagaki, H.; Yamada, K.; Ohye, T.; Taniguchi, M.; Emanuel, B.S.; Toda, T. Cruciform DNA Structure Underlies the Etiology for Palindrome-Mediated Human Chromosomal Translocations. *J. Biol. Chem.* **2004**, *279*, 35377–35383. [[CrossRef](#)] [[PubMed](#)]