The recurrent atypical e8a2 BCR::ABL1 transcript with insertion of an inverted 55 base pair ABL1 intron 1b sequence: a detailed molecular analysis
Burmeister T, Bullinger L, le Coutre P
Case Report

The recurrent atypical e8a2 BCR::ABL1 transcript with insertion of an inverted 55 base pair ABL1 intron 1b sequence: a detailed molecular analysis

Thomas Burmeister\textsuperscript{a,b}, Lars Bullinger\textsuperscript{a}, Philipp le Coutre\textsuperscript{c}

\textsuperscript{a} Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Campus Virchow, Medizinische Klinik für Hämatologie, Onkologie und Tumorimmunologie, Berlin, Germany
\textsuperscript{b} Labor Berlin Charité – Vivantes, 13353 Berlin, Germany
\textsuperscript{c} Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Campus Mitte, Medizinische Klinik für Hämatologie, Onkologie und Tumorimmunologie, Berlin, Germany

Short title: e8a2 BCR::ABL1 with 55 base pair ABL1 intron 1b insertion

Corresponding Author:
Thomas Burmeister
Medizinische Klinik für Hämatologie, Onkologie und Tumorimmunologie, Campus Virchow
Charité – Universitätsmedizin Berlin
Augustenburger Platz 1
13353 Berlin, Germany
Tel: +49 (0)30450653154
E-mail: thomas.burmeister@charite.de

Number of Tables: 0
Number of Figures: 3
Word count: abstract 144 words, body text: 1864 words
Keywords: BCR::ABL1, chromosomal translocation, atypical transcript, cryptic exon
Abstract
Atypical BCR::ABL1 transcripts are found in approximately 2% of cases of chronic myeloid leukemia. It is important to detect them, since affected patients also benefit from tyrosine kinase inhibitor therapy. In the rare e8a2 atypical BCR::ABL1 transcript two out-of-frame exons are fused, thus, interposed nucleotides are usually found at the fusion site to restore the reading frame. In approximately half of previously reported e8a2 BCR::ABL1 cases an inserted 55 bp sequence homologous to an inverted sequence from ABL1 intron 1b was detected. The generation of this recurrent transcript variant is not obvious. This work describes the molecular analysis of such an e8a2 BCR::ABL1 translocation from a CML patient. The genomic chromosomal breakpoint is identified, and the formation of this transcript is theoretically explained. The clinical course of the patient is reported, and recommendations are provided for the molecular analysis of future e8a2 BCR::ABL1 cases.
Introduction
The chimeric BCR::ABL1 gene is the molecular hallmark of chronic myeloid leukemia (CML). More than 95% of BCR::ABL1-positive CML patients reveal a BCR::ABL1 transcript showing a fusion of either BCR exon 13 or 14 and ABL1 exon 2 (transcripts e13a2 and e14a2), indicating a chromosomal breakpoint in the so-called "major breakpoint cluster region" in BCR. Other "atypical" BCR::ABL1 transcripts with different exon fusions are rarely found. One of these is the e8a2 BCR::ABL1 transcript, first described by How et al. (1999) and Branford et al. (2000),[1, 2] of which only a handful of cases have been reported. The generation of this transcript type is more complex since BCR exon 8 and ABL1 exon 2 are not in the same reading frame, thus, a simple fusion of the two exons should not lead to a functional BCR::ABL1 oncoprotein. In fact, the large majority of previously described e8a2 BCR::ABL1 transcripts showed an insertion of additional nucleotides that were derived from intronic sequences of BCR, BCR or other genes, with only very few (difficult to explain) exceptions.[3, 4] Approximately half of the previously reported e8a2 BCR::ABL1 transcript cases showed a 55 base pair (bp) insertion matching an inverted sequence from ABL1 intron 1b.[2, 5–7] How this intronic inverted sequence is incorporated into the final chimeric mRNA transcript is not obvious and has not been investigated previously.

We describe the case of a patient with CML and an e8a2 BCR::ABL1 transcript with such an inverted interposed sequence. We molecularly identified the chromosomal breakpoint and provide a theoretical explanation for the generation of this transcript variant. Implications for molecular monitoring are discussed.

Case Presentation
Patient's history
The patient, a 74-year-old male from Germany presented with a myeloproliferative neoplasm. He presented with a leukocyte count of $110 \times 10^9/L$, hemoglobin 15.6 g/dl and $178 \times 10^9/L$ platelets. The differential blood count showed 5% lymphocytes, 2% eosinophils, 3% basophils, 8% neutrophilic band forms, 57% segmented neutrophils, 10% metamyelocytes, 7% myelocytes, 5% promyelocytes, and 3% blasts. Clinically he was largely asymptomatic. While the karyotype 46,XY,t(9;22)(q34;q11.2)[3] suggested a Philadelphia translocation the initial RT–PCR analysis in an external laboratory failed to detect a BCR::ABL1 transcript. RT–PCR analysis by multiplex PCR in our laboratory and subsequent sequencing of the PCR product revealed an e8a2 BCR::ABL1 transcript with the aforementioned 55 bp insertion.[8] Bone marrow histology showed a myeloproliferative neoplasm with less than 2% blasts and a focal discrete fibrosis. The patient first received hydroxyurea for a few days and then 2×300 mg nilotinib daily. Molecular analysis by RT–qPCR using a quantitative PCR method with a newly designed primer in BCR exon 8 together with the standardized EAC cycling conditions and ABL1 as a housekeeping gene[9, 10] showed a reduction in the relative BCR::ABL1 level to a value of $10^{-3}$ in the peripheral blood after twelve months of therapy (Fig. 1).

Genetic analysis of the BCR::ABL1 break region
The procedures of the genetic analysis are illustrated in Figure 2. Technical details such as PCR methods and primer sequences are provided in the online supplement to this report. All nucleotide positions refer to the human genome assembly GRCh38.p14 139608 bp ABL1 intron 1 (NC_000009.12:130714456-130854063) and 10202 bp BCR intron 8 (NC_000022.11:23273775-23283976). After sequencing the mRNA transcript (Fig. 2a) all further molecular analyses were performed with genomic DNA. The break on chromosome 22 was assumed to be in the 10.2 kb BCR intron 8. Since the inverted intronic sequence was located near ABL1 exon 1b (Fig. 2b), we hypothesized that one ABL1 break occurred near the 5' end of ABL1 exon 1b. Long-range PCRs with different PCR primers located in the inverted sequence and in various parts of BCR exon/intron 8 were performed which resulted in the identification of the 5' break in BCR intron 8 at nt 7894 and in ABL1 intron 1b at nt 1072 (Fig. 2c). To obtain an estimate of the location of the second (3') break, i.e., the size of the inverted segment, various long-range PCRs were performed with one primer located in BCR intron 8 immediately 5' of the breakpoint and the second primer in (inverted) ABL1 at various
distances 3' to the inverted fragment. This revealed that a second break was located still in ABL1 intron 1b and not further 5' upstream.

This 3' break in ABL1 intron 1b was identified using a long-range inverse PCR approach with one primer located in the inverted fragment and the other in BCR intron 8 as shown in Fig. 2c. Three different restriction enzymes were tested and the enzyme TaqI led to the identification of the 3' break in ABL1 intron 1b (Fig. 2c).

The 3' break in ABL1 was located at nt 1415 of intron 1b. An 878 bp sequence fragment spanning from nt 195 to nt 1072 with 100% sequence identity to the ABL1 reference sequence was inserted in reverse complement, while 342 bp of DNA sequence between nt 1073 and 1414 of ABL1 intron 1b were missing (Fig. 3a).

Genetic analysis of the ABL1::BCR breakpoint region
No expression of a chimeric ABL1::BCR transcript was detected by RT–PCR using PCR primers located in ABL1 exon 1b and BCR exon 9.

Bioinformatic analysis of the breakpoint region
The complete breakpoint region in ABL1 and BCR was sequenced from a remission sample of the patient to analyze possible sequence variations that could have facilitated the chromosomal break event. Genomic repeats were analyzed with RepeatMasker version 4.0.9, RSSsite and the Tandem repeats finder.[11–13] No repetitive element was identified in the vicinity of the breakpoint in BCR while four elements were identified in the first 1500 bp of ABL1 intron 1b (elements with intronic location: MIR3 218–322 bp, Charlie4z 400–452 bp, L1ME3G 1226–1333 bp, Charlie1a 1345–1501 bp). No tandem repeats or microhomologies were detected at the break sites. There was one potential 12 bp cryptic recombination signal sequence (cRSS) at 279–306 bp and two potential 23 bp cRSSs at 385–423 bp and at 1242–1280 bp. All three cRSSs were on the (-) strand. No single nucleotide polymorphisms were detected.

Discussion
Various e8a2 BCR::ABL1 transcript variants have been reported in the literature. Many of them were singular cases showing interposed short nucleotide sequences derived from different chromosomes and genes.[14–16] The 55 bp ABL1 insert e8a2 BCR::ABL1 variant is of particular theoretical interest since it has been recurrently observed. None of the previously reported e8a2 BCR::ABL1 cases was analyzed on the genomic level.

Genetic analysis
Previous work has located BCR::ABL1 chromosomal breaks in the first 1.5 kb of ABL1 intron 1b. Linhartova et al. analyzed the der(22) chromosomal breaks of CML patients with e13a2/e14a2 transcripts.[17] Three of the analyzed 46 breaks in ABL1 intron 1b were located in this region (with GenBank accession numbers): at nt 866 (KR092020), nt 841 (KR091994), and nt 604 (KR092015). Krumbholz et al. identified the der(22) and/or der(9) break locations in 59 pediatric CML patients.[18] Six of them mapped to the break region at 575 bp (JQ425201), 698 bp (JQ425193), 1154 bp (JQ425167), 1276 bp (JQ425228), 1280 bp (JQ425229), and 1435 bp (JQ425168). In the author’s previous work however, none of the 64 identified der(22) and der(9) chromosomal breaks of e13a2/e14a2-BCR::ABL1-positive patients were located in this region.[19] Thus there was a higher density of BCR::ABL1 break events in this region than could statistically be expected. When analyzing the breakpoint junctions in the current case no microhomologies, repetitive DNA elements or cryptic recombination signal sequence (RSS) sites were found that were likely involved in the chromosomal break. The question of whether those elements play a role in the formation of the BCR::ABL1 translocation has been controversially discussed.[17–22]

A reciprocal der(9) or ABL1::BCR fusion breakpoint could not be characterized. Since no expression of an ABL1::BCR mRNA transcript was detectable by RT–PCR, it is possible that the translocation led to a partial or complete deletion of the reciprocal allele. Such deletions of der(9) are frequently observed.
in CML[23] and the abovementioned studies also did not detect reciprocal $ABL1::BCR$ breakpoints in a significant percentage of cases.

**Activation of a cryptic exon**
The incorporation of the 55 bp sequence between $BCR$ exon 8 and $ABL1$ exon 2 in the detected mRNA transcript can be theoretically explained by the activation of a cryptic exon (Fig. 3). This has already been hypothesized by Tchirkov et al. (2006) but the authors did not investigate it on the genomic level.[24] The 55 bp cryptic exon possessed intact splice donor and splice acceptor sites, a small polypyrimidine tract and putative splice branching site, which however did not fully match the consensus YNYTRAY motif.[25] Whether this cryptic exon was activated by some kind of abnormality or variant in the splicing apparatus of our patient remains unclear, since there is no known transcriptional unit on the minus strand of the $ABL1$ gene locus, thus, it cannot be decided if this exon is spliced in healthy individuals. As can be seen from the sequence chromatogram in Fig. 2a, a transcript with direct fusion of $BCR$ exon 8 and $ABL1$ exon 2 is also transcribed with low efficiency, but this should not result in a functional $BCR::ABL1$ protein.
The cryptic exon was also predicted with some probability by bioinformatic web-based tools, albeit with only moderate probability: NetGene2-2.42 ($p=0.83$ donor, $p=0.77$ acceptor splice site)[26] and GENSCAN ($p=0.304$).[27]

**Molecular monitoring**
Molecular monitoring of atypical $BCR::ABL1$ transcripts poses a diagnostic problem, since no international guidelines exist for this process. In the case of the typical e13a2/e14a2 $BCR::ABL1$ transcripts the EAC PCR primers are widely used, and detailed recommendations for molecular monitoring have been issued,[28] but no such guidelines exist for atypical transcripts. Recently, Schäfer et al. (2021) addressed this problem and designed and tested PCR primers for several atypical $BCR::ABL1$ transcripts, including e8a2, using a relative patient-specific scale, i.e., the $BCR::ABL1$ transcript level at diagnosis was set to 100% and the relative reduction was calculated during the time course.[29] We adopted this approach, albeit with different PCR primers and probe and our patient showed a significant reduction in the relative $BCR::ABL1$ level to about below $10^{-3}$ (0.1%) after nine months of therapy.

**Conclusions**
This report investigated and elucidated the molecular background of the recurrent e8a2 $BCR::ABL1$ transcript variant with an interposed 55 bp inverted $ABL1$ sequence. The generation of this $BCR::ABL1$ variant is explained by a complex chromosomal translocation involving three chromosomal breaks with partial inversion of and deletion in the $ABL1$ gene locus and subsequent activation of a cryptic exon on the minus DNA strand. Although this transcript variant is recurrent, i.e., repeatedly experimentally observed, it does not imply that the underlying chromosomal break events are identical in different 55 bp insert e8a2 $BCR::ABL1$ cases. To obtain a better understanding of the generation of this variant we suggest the molecular investigation of future e8a2 cases also at the genomic level. The steps undertaken in our work could serve as a model for such an analysis.
Statements
Acknowledgement
The authors wish to thank Ms. Daniela Gröger for her skillful technical work and Ms. Thekla Schwarzer for organizational assistance.

Statement of Ethics
Our research was conducted ethically in accordance with the World Medical Association Declaration of Helsinki. Written informed consent was obtained from the patient for publication of this case report and any accompanying images. This retrospective review of patient data did not require ethical approval in accordance with local and national guidelines.

Conflict of Interest Statement
TB received speakers’ honoraria from Novartis and Pfizer. LB received honoraria from Sanofi, Astellas, Amgen, consultancy fee from Gilead, Hexal, and Menarini, consultancy fee and honoraria from Abbvie, BMS/Celgene, Daiichi Sankyo, Janssen, Jazz Pharmaceuticals, Novartis and Pfizer, and research funding from Bayer and Jazz Pharmaceuticals. PC received speakers’ honoraria from Novartis, BMS, Incyte and Pfizer.

Funding Sources
The work was funded by institutional research support.

Author Contributions
TB designed research and drafted the manuscript. LB and PC treated the patient and contributed to the manuscript. All authors read and approved the final manuscript.

Data Availability Statement
The nucleotide sequence of the chromosomal break sequence has been submitted to the GenBank/EMBL/DDBJ database and is available under accession number OP797408. Further enquiries can be directed to the corresponding author.
References
**Figure Legends**

**Fig. 1.** Time course of the relative *BCR::ABL1* level during twelve months of therapy.

**Fig. 2.** e8a2 transcript and the identification of breaks on chromosomes 9 and 22.
   a: Sequence chromatogram of the *BCR::ABL1* e8a2 transcript.
   b: Exon organization of *BCR* and *ABL1* with the location of the 55 bp insert in *ABL1* intron 1b.
   c. PCR and sequencing strategy to identify the breaks in *BCR::ABL1*. Upper row: long-range PCR for identification of the 5' break. Lower row: long-range inverse PCR for identification of the 3' break(s). Arrows indicate PCR primer locations and orientations. Gel images show PCR products.

**Fig. 3.** Hypothetical mechanism leading to the generation of the observed transcript.
   a. Break region in *ABL1* intron 1b with the three break positions and formation of the *BCR::ABL1* fusion. The position of repetitive DNA elements and cryptic recombination signal sequences are indicated as arrows or bars.
   b. Nucleotide sequence of the cryptic exon with adjacent regions.

Please note that the actual figures and all tables should be uploaded as separate items in their original file format.