

Universal Multiple-Octet Coded Character Set  
International Organization for Standardization  
Organisation Internationale de Normalisation  
Международная организация по стандартизации

**Doc Type:** Working Group Document

**Title:** Proposal to encode one Tangut ideograph

**Source:** Andrew West and Viacheslav Zaytsev

**Status:** Individual Contribution

**Action:** For consideration by JTC1/SC2/WG2 and UTC

**Date:** 2025-03-28

## 1. Introduction

We recently carried out a comprehensive analysis all 6,150 entries in the nine-volume dictionary of Tangut words, *Xīxiàwén Cídiǎn* (*shìsú wénxiàn bùfēn*) 西夏文词典（世俗文献部分）(Běijīng: Zhōngguó shèhuì kēxué chūbǎnshè, 2021), compiled by Hán Xiǎománg 韓小忙. This important new dictionary has already been the source for several submissions to WG2 and the UTC:

- *Proposal to encode 113 Tangut components* by West ([WG2 N5218R3 = L2/23-150](#))
- *Glyph corrections for four Tangut ideographs* by West ([WG2 N5237 = L2/23-148](#))
- *Proposal to encode two Tangut ideographs* by Chan, You, and Zaytsev ([WG2 N5286 = L2/24-236](#))

The Tangut ideographs, Tangut components, and glyph changes proposed in these three documents have been incorporated into Unicode version 17.0.

In this and one other related document (WG2 N5305: *Glyph changes for 14 Tangut ideographs*) we aim to document and resolve any outstanding discrepancies between the Tangut character set in Hán Xiǎománg's dictionary (hereafter *XXWCD*) and the set of Tangut ideographs encoded in the ISO/IEC 10646 and Unicode standards. In this document we document cases where *XXWCD* has two or three entries for glyph variants corresponding to a single encoded Tangut ideograph, and propose to encode one Tangut ideograph which we do not think should be treated as a glyph variant.

## 2. Discussion

There are 147 entries in *XXWCD* which are unencoded variant forms of another entry in the dictionary, and a further six entries for variant glyph forms which are coincidentally identical to the glyph form of a different unrelated Tangut ideograph (see entries for 17009, 17230,

1774A, 177DE, 17AF5, and 187CD in the Appendix).<sup>1</sup> In total, 295 entries in *XXWCD* correspond to 148 Unicode characters (143 Unicode characters correspond to two entries in *XXWCD*, and five Unicode characters correspond to three entries in *XXWCD*). For example, *XXWCD* has entries for 𐺖 (v. 1 p. 39) and 𐺖 (v.1 p. 40), which are glyph variants of the Tangut ideograph meaning ‘lion’ which is encoded as U+17E5B 𐺖 (the Unicode glyph corresponds to the entry on v. 1 p. 40). In most cases the difference between the variants is minor, and either involves a difference in writing a single component (as in the above example) or a difference in positioning of components (e.g. entries for 𐺖, 𐺖 and 𐺖 corresponding to U+177D4 𐺖; and entries for 𐺖 and 𐺖 corresponding to U+182C6 𐺖). One particularly common variation is that between Component 26 𐺖, Component 28 𐺖, and Component 75 𐺖, which are often written in a very similar manner in primary Tangut sources, with the result that *XXWCD* has three entries (𐺖, 𐺖, and 𐺖) corresponding to U+17219 𐺖.

For reference, all the *XXWCD* entries for variants corresponding to the same Unicode character are listed in the Appendix at the end of this document. Almost all of these variant forms can probably be considered to be unifiable glyph variants, and could potentially be represented using Ideographic Variation Sequences (IVS).<sup>2</sup>

However, in one case only we think that the difference between two entries in *XXWCD* is not that of simple glyph variation of the same character. The entry for 𐺖 on v. 5 p.96 (see Fig. 1) defines 𐺖 as a variant form of 𐺖 (v. 4 p. 404), i.e. U+17A11 𐺖, but with a different structure (“异构字”).

𐺖	VII 切韵 2. 20 tsiaa [ ? ]	左 𐺖 / 右 𐺖
𐺖	14 画 1842/22	K. ? L. ? S. ?

𐺖之异体字（异构字）

『异体』参见：切韵 戊 R23。

★按：此字其实与“𐺖”的音值相等，只是反切上字不同而已。异构字“𐺖”的反切上字换成了“𐺖”（tsji）字。






**Fig. 1: *XXWCD* vol. 5 p. 96**

<sup>1</sup> An example of a glyph variant for one character corresponding to the glyph form for another character is the entry for 𐺖 (v. 8 p. 469) which is a variant of U+1774A 𐺖 ‘roof beam’ (𐺖 on v. 3 p. 142), but it is identical to the glyph for U+18736 𐺖 ‘big’ (𐺖 on v. 8 p. 456).

<sup>2</sup> Version 17.000 of the [Tangut Yinchuan](#) font maintained by West includes a provisional set of nearly 3,000 IVS for 2,319 Tangut ideographs, including all the glyph forms used in *XXWCD*. This set may be considered for registration with the [Ideographic Variation Database](#) (IVD) at a future date.

Both characters only occur in manuscript versions of the phonetic text *Joined Rimes of the Five Sounds* 𪛗𪛘𪛙𪛚𪛛 (Chinese 五音切韻 *Wǔyīn Qièyùn*), where they both represent the phonetic syllable *tśiq*<sup>2</sup> in the table for Rime 23 (see Table 1).

**Table 1: Glyphs for *tśiq*<sup>2</sup> in Manuscripts of *Joined Rimes of the Five Sounds***

MS A (甲)	MS B (乙)	MS D (丁)	MS E (戊)	MS F (己)
				
𪛗	𪛗	𪛗	𪛗	𪛗?

Manuscripts A and B show the character 𪛗, Manuscript D shows the variant form 𪛗,<sup>3</sup> and Manuscript F shows an unclear variant of 𪛗, but Manuscript E shows the character 𪛗. 𪛗 and 𪛗 have different left side components (Component 141 𪛗 *versus* Component 210 𪛗), and the two characters appear to have different character constructions. According to *Lǐ Fànwén* 李範文 (*Xià-Hàn Zìdiǎn* 夏漢字典 [2008] p. 949 Entry 6044), 𪛗 is constructed from the left side of 𪛗 *tśiq*<sup>2</sup> and the top/right parts of 𪛗 *śiq*<sup>2</sup> (representing the *fanqie* pronunciation *tśiq*<sup>2</sup>); whereas according to *Hán Xiǎománg* (*XXWCD* v. 5 p. 96), 𪛗 is constructed from the left side of 𪛗 *tśji*<sup>2</sup> and the top/right parts of 𪛗 *śiq*<sup>2</sup> (representing the *fanqie* pronunciation *tśiq*<sup>2</sup>). Although the two characters have the same *fanqie* pronunciation, they are constructed from different characters, and thus should be considered to be *alternative characters* for the syllable *tśiq*<sup>2</sup>, not glyph variants of the same character. Therefore, we propose to encode 𪛗 as a separate character (see Table 2).

**Table 2: Proposed New Tangut Ideograph**

Code Point	Glyph	IDS	Radical/Stroke	Source Reference	XXWCD Vol./Page	XXWCD Glyph
*18D1F	𪛗	𪛗 𪛗 𪛗 𪛗 𪛗 𪛗	210.14	H2021-509604	5/96	𪛗

<sup>3</sup> This variant is included in *XXWCD* as 𪛗 on v. 2 p. 300, and as it is a simple repositioning of the top right component to cover the left component it can be considered to be a unifiable glyph variant of U+17A11.

### 3. Code Point Allocation and Character Properties

#### 3.1. Code Point Allocation

To encode new (recently identified) Tangut ideographs, there are free code points only in the Tangut Supplement block (U+18D00–U+18D7F). The last occupied code point in this block is U+18D1E. This code point was occupied by the Tangut ideograph, which was proposed to be encoded in Proposal WG2 N5286 = L2/24-236 (see References). In this regard, we propose to encode the new character using the following free code point \*U+18D1F.

#### 3.2. Unicode Character Properties

The entries for Tangut ideographs in the UnicodeData.txt file should be amended to:

```
18D00;<Tangut Ideograph Supplement, First>;Lo;0;L;;;;;N;;;;;  
18D1F;<Tangut Ideograph Supplement, Last>;Lo;0;L;;;;;N;;;;;
```

#### 3.3. Tangut Sources

The following entries should be added to the TangutSources.txt file:

```
U+18D1F    kTGT_MergedSrc  H2021-509604  
U+18D1F    kRSTUnicode    210.14
```

### 4. References

#### 4.1. Technical Documents

**N5218R3.** Andrew West. *Proposal to encode 113 Tangut components*. JTC1/SC2/WG2 N5218R3. 2024-06-12.

URL: <https://www.unicode.org/wg2/docs/n5218r3-Tangut-components.pdf>

**N5218R.** Andrew West. *Proposal to encode 114 Tangut components*. JTC1/SC2/WG2 N5218R = L2/23-247. 2023-10-02.

URL: <https://www.unicode.org/wg2/docs/n5218r2-Tangut-components.pdf>

URL: <https://www.unicode.org/L2/L2023/23247-wg2-n5218r2.pdf>

**N5218.** Andrew West. *Proposal to encode 114 Tangut components*. JTC1/SC2/WG2 N5218 = L2/23-150. 2023-05-31.

URL: <https://www.unicode.org/wg2/docs/n5218-Tangut-components.pdf>

URL: <https://www.unicode.org/L2/L2023/23150-n5218-tangut-components.pdf>

**N5237.** Andrew West. *Glyph corrections for four Tangut ideographs*. JTC1/SC2/WG2 N5218 = L2/23-148. 2023-06-21.

URL: <https://www.unicode.org/wg2/docs/n5237-Tangut-glyph-corrections.pdf>

URL: <https://www.unicode.org/L2/L2023/23148-n5237-tangut-glyph-corr.pdf>

**N5286.** Eiso Chan (陈永聪/𐰇𐰆𐰏𐰐), Jerry You (游程宇/𐰇𐰆𐰏𐰐), Viacheslav Zaytsev (IOM RAS). *Proposal to encode two Tangut ideographs*. ISO/IEC JTC1/SC2/WG2 N5286 = L2/24-236. 2024-10-14.

URL: [https://www.unicode.org/wg2/docs/n5286\\_TwoTangutIdeographs.pdf](https://www.unicode.org/wg2/docs/n5286_TwoTangutIdeographs.pdf)

URL: <https://www.unicode.org/L2/L2024/24236-n5286-two-tangut-ideographs.pdf>

## 4.2. Tangut Sources

**Joined Rimes of the Five Sounds.** *ngwə ·êi we mbju* 𐰇𐰆𐰏𐰐𐰏 [Wǔyīn Qièyùn 五音切韻]. Institute of Oriental Manuscripts of the Russian Academy of Sciences, Tangut Collection. MS A (甲): Pressmark Tang 22/1, old inventory no. 620. Published in: ECHW 7, p. 258–278; MS B (乙): Pressmark Tang 22/2, old inventory no. 621. Published in: ECHW 7, p. 279–293; MS C (丙): Pressmark Tang 22/3, old inventory no. 622. Published in: ECHW 7, p. 293–307; MS D (丁): Pressmark Tang 22/4, old inventory no. 623. Published in: ECHW 7, p. 308–368; MS E (戊): Pressmark Tang 22/5, old inventory no. 624. Published in: ECHW 7, p. 369–385; MS F (己): Pressmark Tang 22/10, old inventory no. 7192. Published in: ECHW 7, p. 385–398.

## 4.3. Academic Literature

**ECHW 7.** Shǐ Jīnbō 史金波, Wèi Tóngxián 魏同賢, and E. I. Kychanov (chief eds.). *É cáng Hēishuǐchéng wénxiàn* 俄藏黑水城文獻 = Памятники письменности из Хара-Хото хранящиеся в России = *Heishuicheng Manuscripts Collected in Russia*. Vol. 7. Shànghǎi 上海: Shànghǎi gǔjí chūbǎnshè 上海古籍出版社, 1997. ISBN 7-5325-2213-X.

**Xià-Hàn Zìdiǎn.** Lǐ Fànwén 李範文 (comp.) and Jiǎ Chángyè 賈常業 (rev. and exp.). *Xià-Hàn zìdiǎn* 夏漢字典 [Tangut-Chinese Dictionary (Revised Ed.)]. Běijīng 北京: Zhōngguó shèhuì kēxué chūbǎnshè 中國社會科學出版社, 2008. ISBN 978-7-5004-2113-9.

**XXWCD.** Hán Xiǎománg 韩小忙 (comp.). *Xīxiàwén cídiǎn* (shìsú wénxiàn bùfēn) 西夏文词典 (世俗文献部分): 全 9 册 [Tangut Word Dictionary (Secular Literature Part). 9 vols.]. Běijīng 北京: Zhōngguó shèhuì kēxué chūbǎnshè 中国社会科学出版社, 2021. (*Xīxià wénzì yǔ wénxiàn yánjiū* 西夏文字与文献研究 [Studies of Tangut characters and documents]). ISBN 978-7-5203-8011-9.

**Appendix: XXWCD Entries for Variants of the Same Unicode Character**

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17009	𣦵	1.8	1/10	𣦵	LR1.8
			1/17 <sup>4</sup>	𣦵	LR1.9
1708D	𣦵	2.12	1/99	𣦵	LR2.12
			1/100	𣦵	LR2.12
170BD	𣦵	2.14	1/143	𣦵	LR2.14
			1/133	𣦵	LR2.13
170F0	𣦵	9.10	1/174	𣦵	LR4.10
			2/593	𣦵	LR47.10
1710E	𣦵	11.11	1/192	𣦵	LR5.11
			1/200	𣦵	LR5.11
17118	𣦵	11.14	1/197	𣦵	LR5.14
			1/197	𣦵	LR5.14

<sup>4</sup> This variant of U+17009 is identical to U+17011 𣦵 (𣦵 on v. 1 p. 17).

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17150	𠂔	17.9	1/331	𠂔	LR7.9
			1/331	𠂔	LR7.9
17155	𠂕	17.10	1/335	𠂕	LR7.10
			1/335	𠂕	LR7.10
171A6	𠂖	17.13	1/391	𠂖	LR7.13
			1/391	𠂖	LR7.13
171AD	𠂗	17.14	1/398	𠂗	LR7.14
			1/398	𠂗	LR7.14
171B0	𠂘	17.14	1/402	𠂘	LR7.14
			1/403	𠂘	LR7.14
171BD	𠂙	17.16	1/409	𠂙	LR7.16
			1/409	𠂙	LR7.16
171D8	𠂚	26.13	1/430	𠂚	LR10.13
			1/462	𠂚	LR12.13

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
171E5	𪛥	28.9	1/438	𪛥	LR12.9
			2/235	𪛦	LR36.10
171EA	𪛧	28.10	1/440	𪛧	LR12.10
			1/424	𪛨	LR10.10
171EB	𪛩	28.10	1/441	𪛩	LR12.10
			2/252	𪛪	LR36.11
171F7	𪛫	28.11	1/448	𪛫	LR12.11
			2/274	𪛬	LR36.12
171F9	𪛭	28.11	1/449	𪛭	LR12.11
			1/428	𪛮	LR10.11
171FC	𪛯	28.11	1/450	𪛯	LR12.11
			2/277	𪛰	LR36.12
171FF	𪛱	28.12	1/451	𪛱	LR12.12
			1/455	𪛲	LR12.12



Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
1720B	𪛗	28.13	1/458	𪛗	LR12.13
			1/429	𪛗	LR10.13
1720C	𪛘	28.13	1/460	𪛘	LR12.13
			1/465	𪛘	LR12.13
1720D	𪛚	28.13	1/460	𪛚	LR12.13
			1/461	𪛚	LR12.13
17212	𪛜	28.13	1/465	𪛜	LR12.13
			2/346	𪛜	LR36.14
17219	𪛞	28.14	1/466	𪛞	LR12.14
			1/431	𪛞	LR10.14
			2/357	𪛞	LR36.15
1721A	𪛠	28.14	1/468	𪛠	LR12.14
			2/358	𪛠	LR36.15

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
1721B	𪛮	28.14	1/468	𪛮	LR12.14
			2/359	𪛮	LR36.15
1721C	𪛯	28.14	1/468	𪛯	LR12.14
			1/431	𪛯	LR10.14
			2/359	𪛯	LR36.15
1721F	𪛰	28.15	1/472	𪛰	LR12.15
			1/432	𪛰	LR10.15
17223	𪛱	28.16	1/474	𪛱	LR12.16
			2/388	𪛱	LR36.17
17224	𪛲	28.17	1/474	𪛲	LR12.17
			1/474	𪛲	LR12.17
			2/392	𪛲	LR36.18
17228	𪛳	28.19	1/477	𪛳	LR12.19
			1/477	𪛳	LR12.19

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17230	𢉨	31.8	1/498	𢉨	LR13.8
			3/483 <sup>5</sup>	𢉨	LR88.10
17241	𢉩	33.9	1/516	𢉩	LR14.9
			1/516	𢉩	LR14.9
17289	𢉪	36.7	1/566	𢉪	LR16.7
			1/565	𢉪	LR16.7
172F6	𢉫	36.12	1/644	𢉫	LR16.12
			1/644	𢉫	LR16.12
17311	𢉬	36.14	1/671	𢉬	LR16.14
			1/671	𢉬	LR16.14
1731C	𢉭	36.14	1/675	𢉭	LR16.14
			1/675	𢉭	LR16.14
1733C	𢉮	41.9	2/12	𢉮	LR19.9
			2/14	𢉮	LR19.10

<sup>5</sup> This variant of U+17230 is identical to U+17BE2 𢉨 (𢉨 on v. 3 p. 482).

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17382	𠄎	47.13	2/79	𠄎	LR23.13
			2/78	𠄎	LR23.12
173D3	𠄎	68.9	2/159	𠄎	LR32.9
			2/163	𠄎	LR32.10
17442	𠄎	75.10	2/224	𠄎	LR36.10
			1/438	𠄎	LR12.9
1746B	𠄎	75.11	2/262	𠄎	LR36.11
			2/288	𠄎	LR36.12
174A9	𠄎	75.13	2/294	𠄎	LR36.13
			8/127	𠄎	LR303.13
174B0	𠄎	75.13	2/297	𠄎	LR36.13
			2/297	𠄎	LR36.13
174DC	𠄎	75.14	2/334	𠄎	LR36.14
			2/337	𠄎	LR36.14

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
174E1	𪛑	75.14	2/335	𪛑	LR36.14
			2/352	𪛑	LR36.15
17540	𪛒	76.8	2/403	𪛒	LR37.8
			2/404	𪛒	LR37.8
1754E	𪛓	79.9	2/411	𪛓	LR39.9
			2/450	𪛓	LR39.9
17574	𪛔	79.15	2/449	𪛔	LR39.15
			6/299	𪛔	LR169.15
1762B	𪛕	92.10	2/593	𪛕	LR47.10
			8/445	𪛕	LR390.10
17643	𪛖	99.10	3/11	𪛖	LR50.10
			3/12	𪛖	LR50.10
176AD	𪛗	106.11	3/64	𪛗	LR55.11
			3/64	𪛗	LR55.11

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
176B8	蔞	106.11	3/67	蔞	LR55.11
			3/212	蔞	LR55.11
1771D	蔞	106.14	3/124	蔞	LR55.14
			3/192	蔞	LR55.17
1773E	蔞	106.14	3/138	蔞	LR55.14
			3/138	蔞	LR55.14
1774A	蔞	106.14	3/142	蔞	LR55.14
			8/469 <sup>6</sup>	蔞	LR395.14
1774E	蔞	106.14	3/143	蔞	LR55.14
			3/119	蔞	LR55.13
177AF	蔞	106.17	3/196	蔞	LR55.17
			3/196	蔞	LR55.17
177BB	蔞	106.18	3/199	蔞	LR55.18
			3/199	蔞	LR55.18

<sup>6</sup> This variant of U+1774A is identical to U+18736 蔞 (蔞 on v. 8 p. 456).

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
177D4	𪗇	106.21	3/208	𪗇	LR55.21
			3/208	𪗇	LR55.21
			3/211	𪗇	LR55.21
177DE	𪗈	107.16	8/34	𪗈	LR284.16
			8/577 <sup>7</sup>	𪗈	LR435.16
177EA	𪗉	108.11	3/225	𪗉	LR56.11
			3/227	𪗉	LR56.12
17850	𪗊	112.15	3/289	𪗊	LR58.15
			3/289	𪗊	LR58.15
17890	𪗋	138.9	3/350	𪗋	LR68.9
			3/369	𪗋	LR74.9
17891	𪗌	138.9	3/350	𪗌	LR68.9
			3/369	𪗌	LR74.9

<sup>7</sup> This variant of U+177DE is identical to U+187CD 𪗈 (𪗈 on v. 8 p. 576).

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17892	𠄂	138.9	3/350	𠄂	LR68.9
			3/369	𠄃	LR74.9
17893	𠄄	138.10	3/350	𠄄	LR68.10
			3/369	𠄅	LR74.10
17894	𠄆	138.10	3/351	𠄆	LR68.10
			3/369	𠄇	LR74.10
17896	𠄈	138.10	3/354	𠄈	LR68.10
			3/370	𠄉	LR74.10
17897	𠄊	138.11	3/355	𠄊	LR68.11
			3/370	𠄋	LR74.11
17899	𠄌	138.11	3/356	𠄌	LR68.11
			3/370	𠄍	LR74.11
1789C	𠄎	138.11	3/358	𠄎	LR68.11
			3/370	𠄏	LR74.11



Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
1789E	𠄎	138.12	3/358	𠄎	LR68.12
			3/370	𠄎	LR74.12
1789F	𠄏	138.12	3/359	𠄏	LR68.12
			3/370	𠄏	LR74.12
178A3	𠄐	138.13	3/360	𠄐	LR68.13
			3/370	𠄐	LR74.13
17916	𠄑	141.10	4/123	𠄑	LR70.10
			4/217	𠄑	LR70.11
17982	𠄒	141.11	4/264	𠄒	LR70.11
			4/614	𠄒	LR71.11
17986	𠄓	141.11	4/265	𠄓	LR70.11
			4/176	𠄓	LR70.10
179DA	𠄔	141.12	4/360	𠄔	LR70.12
			4/615	𠄔	LR71.12

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
179F7	𪛏	141.13	4/379	𪛏	LR70.13
			4/378	𪛐	LR70.13
179FC	𪛑	141.13	4/386	𪛑	LR70.13
			4/447	𪛒	LR70.14
17A0C	𪛓	141.13	4/400	𪛓	LR70.13
			4/453	𪛔	LR70.14
17A11	𪛕	141.13	4/404	𪛕	LR70.13
			2/300	𪛖	LR36.13
			5/96	𪛗	LR116.14
17A35	𪛘	141.13	4/428	𪛘	LR70.13
			4/362	𪛙	LR70.12
17A36	𪛚	141.13	4/428	𪛚	LR70.13
			4/364	𪛛	LR70.12

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17AC8	𪗇	141.16	4/536	𪗇	LR70.16
			4/536	𪗇	LR70.16
17AF5	𪗊	142.8	4/26 <sup>8</sup>	𪗊	LR70.8
			4/589	𪗊	LR71.8
17B0B	𪗋	142.15	4/616	𪗋	LR71.15
			4/616	𪗋	LR71.15
17B45	𪗍	152.15	3/397	𪗍	LR79.15
			3/397	𪗍	LR79.15
17B4F	𪗏	155.10	3/410	𪗏	LR80.10
			3/413	𪗏	LR80.11
17BD5	𪗑	170.7	3/474	𪗑	LR87.7
			3/482	𪗑	LR88.7
17BD8	𪗒	170.10	3/477	𪗒	LR87.10
			3/482	𪗒	LR88.10

<sup>8</sup> This variant of U+17AF5 is identical to U+178BB 𪗊 (𪗊 on v. 4 p. 26).

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17C6F	𪗇	185.12	3/610	𪗇	LR99.12
			3/609	𪗇	LR99.12
17CA9	𪗉	185.16	6/410	𪗉	LR177.16
			6/411	𪗉	LR177.16
17CAB	𪗊	185.16	3/648	𪗊	LR99.16
			3/647	𪗊	LR99.16
17D37	𪗋	206.13	5/9	𪗋	LR112.13
			8/559	𪗋	LR428.13
17D42	𪗌	206.14	5/23	𪗌	LR112.14
			5/48	𪗌	LR112.15
17DA1	𪗍	216.13	6/230	𪗍	LR163.13
			8/551	𪗍	LR422.13
17DC8	𪗎	217.15	5/129	𪗎	LR119.15
			5/129	𪗎	LR119.15

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
17E5B	𪛻	240.13	1/40	𪛻	LR1.13
			1/39	𪛻	LR1.13
17E65	𪛽	243.13	5/257	𪛽	LR132.13
			5/257	𪛽	LR132.13
17E87	𪛿	256.9	5/296	𪛿	LR137.9
			6/409	𪛿	LR176.9
17E88	𪛿	256.11	5/296	𪛿	LR137.11
			6/409	𪛿	LR176.11
17EC3	𪛻	260.12	5/324	𪛻	LR139.12
			5/316	𪛻	LR139.11
17EEA	𪛻	260.14	5/334	𪛻	LR139.14
			5/334	𪛻	LR139.14
17F85	𪛻	264.16	5/397	𪛻	LR142.16
			5/397	𪛻	LR142.16

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
18035	𪗇	278.13	6/129	𪗇	LR149.13
			6/129	𪗈	LR149.13
180B1	𪗉	298.10	6/207	𪗉	LR161.10
			6/212	𪗊	LR161.11
1811B	𪗋	310.17	6/299	𪗋	LR169.17
			2/452	𪗌	LR39.17
18123	𪗍	312.11	6/308	𪗍	LR171.11
			6/309	𪗎	LR171.12
1812A	𪗏	314.9	6/313	𪗏	LR172.9
			6/314	𪗐	LR172.9
1814A	𪗑	316.12	6/357	𪗑	LR173.12
			6/357	𪗒	LR173.12
1816F	𪗓	316.15	6/396	𪗓	LR173.15
			6/397	𪗔	LR173.15

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
181A4	𪛮	328.10	6/434	𪛮	LR181.10
			7/210	𪛮	LR238.11
18207	𪛯	345.10	6/497	𪛯	LR195.10
			1/169	𪛯	LR4.10
182C6	𪛰	401.8	7/161	𪛰	LR232.8
			7/161	𪛰	LR231.8
182C7	𪛱	401.14	7/161	𪛱	LR232.14
			7/161	𪛱	LR231.14
18320	𪛲	416.12	3/429	𪛲	LR83.12
			7/234	𪛲	LR239.12
1834E	𪛳	420.11	7/254	𪛳	LR242.11
			7/243	𪛳	LR241.11
18380	𪛴	431.13	7/297	𪛴	LR248.14
			7/296	𪛴	LR248.13

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
183BA	𪗇	436.10	7/346	𪗇	LR251.10
			7/346	𪗈	LR251.10
18432	𪗉	436.15	7/448	𪗉	LR251.15
			7/449	𪗊	LR251.15
18466	𪗋	446.10	7/476	𪗋	LR256.10
			8/128	𪗌	LR304.11
18467	𪗍	446.13	7/476	𪗍	LR256.13
			8/131	𪗎	LR304.14
18468	𪗏	446.13	7/476	𪗏	LR256.13
			8/131	𪗐	LR304.14
18478	𪗑	452.12	7/489	𪗑	LR261.12
			7/493	𪗒	LR262.12
1847B	𪗓	452.13	7/491	𪗓	LR261.13
			7/493	𪗔	LR262.13



Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
1847F	𠂔	453.11	7/492	𠂔	LR262.11
			7/487	𠂔	LR261.11
18539	𠂕	500.15	8/77	𠂕	LR293.15
			8/77	𠂕	LR293.15
1854B	𠂖	512.11	8/111	𠂖	LR301.11
			8/121	𠂖	LR301.11
1856F	𠂗	526.10	8/143	𠂗	LR309.10
			8/143	𠂗	LR309.10
1867D	𠂘	574.11	8/334	𠂘	LR340.11
			2/593	𠂘	LR47.11
1868F	𠂙	585.17	3/213	𠂙	LR55.17
			3/213	𠂙	LR55.17
186AC	𠂚	600.12	8/354	𠂚	LR354.12
			8/354	𠂚	LR354.12

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
186BB	𪛗	601.16	8/362	𪛗	LR355.16
			8/362	𪛗	LR355.16
1875B	𪛘	684.13	8/492	𪛘	LR399.13
			8/487	𪛘	LR399.13
1875C	𪛙	684.15	8/494	𪛙	LR399.15
			8/487	𪛙	LR399.15
18779	𪛛	696.13	8/517	𪛛	LR406.13
			8/517	𪛛	LR407.13
1877A	𪛜	696.15	8/517	𪛜	LR406.15
			8/517	𪛜	LR407.15
18793	𪛞	708.17	1/202	𪛞	LR5.17
			7/16	𪛞	LR198.17
187C7	𪛟	730.11	8/574	𪛟	LR432.11
			8/575	𪛟	LR433.11

Unicode 17.0 Alpha			XXWCD		
Code Point	Glyph	R/S	Vol./Page	Glyph	R/S
187C8	𪗀	730.17	8/574	𪗀	LR432.17
			8/575	𪗀	LR433.17
187C9	𪗁	730.19	8/575	𪗁	LR432.19
			8/575	𪗁	LR433.19
187CA	𪗂	731.22	3/213	𪗂	LR55.22
			3/213	𪗂	LR55.22
187CD	𪗃	733.16	8/35 <sup>9</sup>	𪗃	LR284.16
			8/576	𪗃	LR435.16
187D6	𪗄	739.15	2/595	𪗄	LR47.15
			2/595	𪗄	LR47.15
187DD	𪗅	744.16	8/593	𪗅	LR440.16
			8/592	𪗅	LR440.16

<sup>9</sup> This variant of U+187CD is identical to U+177DE 𪗃 (𪗃 on v. 8 p. 34).

## Proposal Summary Form

**ISO/IEC JTC 1/SC 2/WG 2**  
**PROPOSAL SUMMARY FORM TO ACCOMPANY SUBMISSIONS**  
**FOR ADDITIONS TO THE REPERTOIRE OF ISO/IEC 10646<sup>10</sup>**

Please fill all the sections A, B and C below.

Please read Principles and Procedures Document (P & P) from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/principles.html> for guidelines and details before filling this form.

Please ensure you are using the latest Form from <http://www.dkuug.dk/JTC1/SC2/WG2/docs/summaryform.html>.

See also <http://www.dkuug.dk/JTC1/SC2/WG2/docs/roadmaps.html> for latest Roadmaps.

### A. Administrative

1. Title:	<b>Proposal to encode one Tangut ideograph</b>		
2. Requester's name:	<b>Andrew West and Viacheslav Zaytsev</b>		
3. Requester type (Member body/Liaison/Individual contribution):	<b>Individual contribution</b>		
4. Submission date:	<b>2025-03-28</b>		
5. Requester's reference (if applicable):			
6. Choose one of the following:			
This is a complete proposal:			<b>YES</b>
(or) More information will be provided later:			

### B. Technical – General

1. Choose one of the following:			
a. This proposal is for a new script (set of characters):			<b>NO</b>
Proposed name of script:			
b. The proposal is for addition of character(s) to an existing block:			<b>YES</b>
Name of the existing block:	<b>Tangut Supplement</b>		
2. Number of characters in proposal:			<b>1</b>
3. Proposed category (select one from below - see section 2.2 of P&P document):			
A-Contemporary	B.1-Specialized (small collection)	B.2-Specialized (large collection)	
C-Major extinct	D-Attested extinct	E-Minor extinct	
F-Archaic Hieroglyphic or Ideographic	<b>X</b>	G-Obscure or questionable usage symbols	
4. Is a repertoire including character names provided?			<b>YES</b>
a. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?			<b>YES</b>
b. Are the character shapes attached in a legible form suitable for review?			<b>YES</b>
5. Fonts related:			
a. Who will provide the appropriate computerized font to the Project Editor of 10646 for publishing the standard?	<b>Andrew West</b>		
b. Identify the party granting a license for use of the font by the editors (include address, e-mail, ftp-site, etc.):	<b>Andrew West</b>		
6. References:			
a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?			<b>YES</b>
b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?			<b>NO</b>
7. Special encoding issues:			
Does the proposal address other aspects of character data processing (if applicable) such as input, presentation, sorting, searching, indexing, transliteration etc. (if yes please enclose information)?			<b>NO</b>

### 8. Additional Information:

Submitters are invited to provide any additional information about Properties of the proposed Character(s) or Script that will assist in correct understanding of and correct linguistic processing of the proposed character(s) or script. Examples of such properties are: Casing information, Numeric information, Currency information, Display behaviour information such as line breaks, widths etc., Combining behaviour, Spacing behaviour, Directional behaviour, Default Collation behaviour, relevance in Mark Up contexts, Compatibility equivalence and other Unicode normalization related information. See the Unicode standard at <http://www.unicode.org> for such information on other scripts. Also see Unicode Character Database ( <http://www.unicode.org/reports/tr44/> ) and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

<sup>10</sup> Form number: N4102-F (Original 1994-10-14; Revised 1995-01, 1995-04, 1996-04, 1996-08, 1999-03, 2001-05, 2001-09, 2003-11, 2005-01, 2005-09, 2005-10, 2007-03, 2008-05, 2009-11, 2011-03, 2012-01)

## C. Technical - Justification

1. Has this proposal for addition of character(s) been submitted before?	NO
If YES explain	
2. Has contact been made to members of the user community (for example: National Body, user groups of the script or characters, other experts, etc.)?	YES
If YES, with whom? Other experts	
If YES, available relevant documents:	
3. Information on the user community for the proposed characters (for example: size, demographics, information technology use, or publishing use) is included?	NO
Reference:	
4. The context of use for the proposed characters (type of use; common or rare)	rare
Reference:	
5. Are the proposed characters in current use by the user community?	YES
If YES, where? Reference:	
6. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?	NO
If YES, is a rationale provided?	
If YES, reference:	
7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?	YES
8. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
9. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
10. Can any of the proposed character(s) be considered to be similar (in appearance or function) to, or could be confused with, an existing character?	NO
If YES, is a rationale for its inclusion provided?	
If YES, reference:	
11. Does the proposal include use of combining characters and/or use of composite sequences?	NO
If YES, is a rationale for such use provided?	
If YES, reference:	
Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?	
If YES, reference:	
12. Does the proposal contain characters with any special properties such as control function or similar semantics?	NO
If YES, describe in detail (include attachment if necessary)	
13. Does the proposal contain any Ideographic compatibility characters?	NO
If YES, are the equivalent corresponding unified ideographic characters identified?	
If YES, reference:	