

2 Comments on some glyphs

code points and position	current	desired	comments
1820 180B second form (medial)			The current glyph on the Unicode 9.0 code chart is a duplicate of the third medial form. The desired glyph is the one in the Unicode 8.0 list of standardized variants and in GB standards.
186F 180B second form (initial)			The short descender in the current glyphs should be removed as it prevents the glyph from joining smoothly with the following letter, which can only be or . The descender is not present in Unicode 8.0 or GB/T 26226-2010. Compare (current) with (desired).
186F 180B second form (medial)			
185E 180C third form (medial)			The current glyph is a duplicate of the first medial form. It does not join smoothly with the preceding letter, which can only be or . The desired glyph is identical to the fourth medial form of U+1873, as given in 《蒙古文编码》 and GB/T 26226-2010. Compare (current) with (desired).
185D first form (isolate)	not shown		These isolate forms need to be shown as they differ from the code point glyphs. The first initial/medial/final forms of these characters were added in the Unicode 9.0 code chart.
185E first form (isolate)			
1873 first form (isolate)			
1869 first form (final)		not shown	These glyphs added in the Unicode 9.0 code chart should be removed as they are not attested in the indicated positions. They are not given in 《蒙古文编码》 or GB standards.
1876 first form (final)			
1887 first form (initial)			
1887 first form (medial)			
188A first form (final)			
1887 first form (isolate)	not shown		The desired glyphs are that of Mongolian and Todo Ali Gali A. The current glyph is the final form of Manchu Ali Gali AH, but the two letters cannot be unified. See the next page.
1887 code point glyph			
1887 first form (final)			
1887 180D fourth form (final)			see below

The fourth final form of U+1887 MONGOLIAN LETTER ALI GALI A is given the glyph in Unicode 8.0, and in Unicode 9.0. Neither glyph should be encoded as a final form of U+1887:

- In Mongolian, these glyphs are precursors in an older style of the backward tail of U+1820 A, U+1821 E, and U+1828 NA, and should be encoded at these code points. Several variants of these glyphs have indeed been proposed as final forms of U+1820 and U+1828 in L2/16-261.

- In Manchu Ali Gali, 𐎡 transcribes Sanskrit long vowel ā and the current glyph is a decorative variant thereof. Either glyph is in fact a sequence of two letters: 𐎠 (medial form of U+1820) followed by 𐎡 (final form of MANCHU ALI GALI AH).

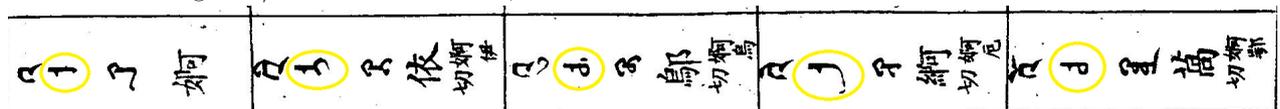
Mongolian/Todo Ali Gali A and Manchu Ali Gali AH. The contextual forms of the two letters are detailed in the following table.

character	𐎡 18AE Manchu Ali Gali AH	𐎠 1887 Mongolian/Todo Ali Gali A
isolate form	𐎡	𐎠
initial form	𐎡	none
medial form	none	none
final form	𐎡	𐎠 first form default
		𐎠 second form context-driven: 18A6 1887 ✓
		𐎠 third form context-driven: 188A 1887 𐎠/𐎡
		𐎠 fourth form post-MVS, limited usage

In Mongolian/Todo, the shape 𐎡 does not represent A, but rather is part of the final forms of several letters. In the example 𐎠 𐎡 *bud dha* from 《大藏全咒》, the shape 𐎡 at the end of the first syllable is part of the final form of U+1891 MONGOLIAN LETTER ALI GALI DA², representing the consonant *d* with no following vowel.

As can be seen from the table above, Mongolian/Todo Ali Gali A and Manchu Ali Gali AH do not share any common glyph. They also have different linguistic values: Mongolian/Todo Ali Gali A is a vowel, while Manchu Ali Gali AH transcribes U+0F60 ᠠ, Tibetan consonant letter -A. For these reasons, the two letters cannot be unified and should be encoded as different characters.

The following table from 《同文韻統》 gives examples of Manchu Ali Gali AH (circled) corresponding to another Mongolian/Todo Ali Gali letter, U+1897 MONGOLIAN LETTER ALI GALI AH.



3 Preliminary proposal to encode four letters in Todo and Manchu Ali Gali

The proposed characters and glyphs are the following:

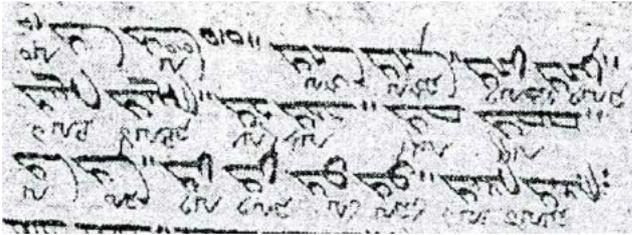
- 18AB 𐎡 MONGOLIAN LETTER TODO ALI GALI REVERSED I
~ 18AB 𐎡 (final)
- 18AC 𐎡 MONGOLIAN LETTER MANCHU ALI GALI VOCALIC R
~ 18AC 𐎡 (initial)
~ 18AC 𐎡 (medial)
- 18AD 𐎡 MONGOLIAN LETTER MANCHU ALI GALI GA
~ 18AD 𐎡 first form (initial)
~ 18AD 𐎡 first form (medial)

²In Baiti, the final form would look like 𐎡.

~ 18AD 180B ᠰ̇ second form (medial)

- 18AE  MONGOLIAN LETTER MANCHU ALI GALI AH
 ~ 18AE  (isolate)
- 18AE  (initial)
- 18AE  (final)

U+18AB Todo Ali Gali Reversed I. This letter transcribes U+0F80 [◌] TIBETAN VOWEL SIGN REVERSED I and is used in the transcription of the Sanskrit vowels ङ and ञ. The glyph is derived by reflection from ᠰ̇, the final form of U+1845 MONGOLIAN LETTER TODO I. The two letters have different sound values. Examples can be found in the following sample of Zaya Bandida's .



On a side note, the above sample also shows that U+1843 [◌] MONGOLIAN LETTER TODO LONG VOWEL SIGN is not a letter, but a diacritic mark.

U+18AC Manchu Ali Gali Vocalic R. This letter is used in the transcription of Sanskrit vowel ɾ. It differs from U+1875 MONGOLIAN LETTER MANCHU RA in shape and sound value, as shown by the following examples from 《同文韻統》:

- 18AC 1873  ɾ
- 1875 1873  ri

Manchu Ali Gali vocalic R also appears in the medial position, as shown in the seventh and eighth rows of the following table from 《同文韻統》.

												
噶	噶	噶	噶	噶	噶	噶	噶	噶	噶	噶	噶	噶
												
ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ	ᠭ

U+18AD Manchu Ali Gali GA. This letter is derived from the masculine form of U+1864 MONGOLIAN LETTER SIBE GA. In contrast to U+1864, U+18AD takes the masculine form even when followed by the vowels i and e. When followed by i and e, the two letters differ in shape and sound value, as shown by the following examples from 《同文韻統》 (see circled items in the table above for context):

18AD	 ᠭᠢ gi	 ᠭᠡ ge
1864	 ᠭᠢ ki	 ᠭᠢ ke

The sequence 18AD 185D *ge* forms a ligature in the isolate, initial, and medial positions. In the final position, 18AD 185D *ge* is attested in two different shapes, making a second medial form of 18AD necessary.

U+18AE Manchu Ali Gali AH. See the last part of Section 2.

4 Proposal to define six variation sequences

Many variants in the Mongolian block were proposed in WG2 N4752R2 (L2-16/258, DS01) and WG2 N4757 (L2-16/261). We propose five additional variants and provide more information on one previously proposed variant.

Very common variants

In Manchu and Sibe words consisting of a consonant (or a sequence of consonants in Manchu Ali Gali) followed by U+1860  MONGOLIAN LETTER SIBE UE, the vowel takes one of the following forms, depending on the preceding consonant.

proposed form	glyph	comments
1860 180C third form (final)		context-driven: used in monosyllabic words when preceded by a consonant other than 1868/1869/18A0/18A1.
1860 180D fourth form (final)		context-driven: used in monosyllabic words when preceded by 1868/1869/18A0/18A1. Examples:  .

The third final form is included in GB/T 26226-2010 and DS01. It is included here because of its high frequency and to provide more information.

Both forms are very commonly used in native words and in transcriptions of other languages. Without either of the two forms, the most basic syllabary (*a e i uju*) cannot be rendered, as shown in the following sample of 《清漢對音字式》.



Other variants

proposed form	glyph	comments	examples
1859 180B second form (medial)		context-driven: used in transcriptions of aspirated voiced consonants	see the Todo sample in Section 3
185D 180C third form (medial)	-	context-driven: 189B 185D (initial)  18AE 185D (initial) 	《同文韻統》  𐩪 ngei  𐩪 ngen
185D 180C third form (final)		context-driven: 189B 185D (isolate)  18AE 185D (isolate)  18AD 185D (final) 	《同文韻統》  𐩪 nge
1865 180B second form (medial)		The default medial/final form of <i>hū</i> is 1865 1861  /  . In a few Manchu words, <i>hū</i> takes the exceptional form 1865 180B 1861  /  .	《五體清文鑑》 