

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 for encoding the Chakma script in the UCS

Source: UC Berkeley Script Encoding Initiative (Universal Scripts Project)

Authors: Michael Everson and Martin Hosken

Status: Liaison Contribution

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

Date: 2009-08-13

Replaces: N3428

1. Introduction. The Chakma people were in origin Tibeto-Burman, related to the Burmese. The language which they now speak is Indo-European, part of the Southeastern Bengali branch of Eastern Indo-Aryan. Its better-known closest relatives are Bengali, Assamese, Chittagonian, Bishnupriya, and Sylheti. It is spoken by 312,000 people in southeast Bangladesh near Chittagong City, and another 176,000 in India in Mizoram, Assam, Tripura, and Arunachal Pradesh. Literacy in Chakma script is low. The script itself is also called ᱡᱷᱟ ᱛᱟᱞᱟ *Ajhā pāṭh*, sometimes romanized *Ojhopath*.

There is a certain amount of glyph variation between the script as used in India and Bangladesh. Some fonts are rounder, similar to the style used in Myanmar; compare a similar variation in the Tai Tham script as used in Khün Tai. The glyphs used in this proposal are based on the Chadigang font, with some alterations toward more “generic” shapes for some characters.

The Chakma script is currently being adapted for use in Tanchangya, a language which is closely related to Chakma. An effort to develop the orthography is currently underway, and it appears that there may be additional letters, vowel signs, and tone marks added to cover this script. These extensions are a subject for future standardization, as the orthography for Tanchangya is still under development and testing.

2. Structure. Chakma is of the Brahmic type: the consonant letters contain an inherent vowel. Consonant clusters are written with conjunct characters, and a visible vowel killer shows the deletion of the inherent vowel when there is no conjunct.

3. Independent vowels. Four independent vowels exist: ᱠ *a*, ᱡ *i*, ᱢ *u*, and ᱣ *e*. Other vowels in initial position are formed by adding the vowel sign to ᱠ *a*, as in ᱠᱢ *ī*, ᱠᱣ *ū*, ᱠᱤ *ai*, ᱠᱥ *oi*. Some modern writers are generalizing this spelling in ᱠᱢ *i*, ᱠᱣ *u*, and ᱠᱤ *e*.

4. Dependent vowels. Independent vowel signs have been encoded according to their phonetic value, as in Balinese or Telugu. However, because some of the independent vowels appear to be made out of smaller units, decomposition of the vowels has been taken into consideration. In order to avoid any ambiguous encoding, canonical equivalences have been given in order to resolve the issue of multiple representations. Two glyph fragments (CHAKMA O MARK and CHAKMA AU MARK) have been encoded in order to account for this, and to allow for any users who desire to create texts using the lower glyph fragments on their own. The Chakma vowel signs are given with the letter ᱠ *ka* below.

𑌓	kā	=	𑌓	kā
𑌔	ka	=	𑌓	kā + 𑌔 -a (11127)
𑌕	ki	=	𑌓	kā + 𑌕 i (11128)
𑌖	kī	=	𑌓	kā + 𑌖 -ī (11129)
𑌗	ku	=	𑌓	kā + 𑌗 -u (1112A)
𑌘	kū	=	𑌓	kā + 𑌘 -ū (1112B)
𑌙	ke	=	𑌓	kā + 𑌙 -e (1112C)
𑌚	kāi	=	𑌓	kā + 𑌚 -āi (1112D)
𑌛	ko	=	𑌓	kā + 𑌛 -o (1112E – or – 11131 + 11127)
𑌜	kau	=	𑌓	kā + 𑌜 -au (1112F – or – 11132 + 11127)
𑌝	koi	=	𑌓	kā + 𑌝 -oi (11130)
𑌞	kaṁ	=	𑌓	kā + 𑌞 -ṁ (11100)
𑌟	kaṃ	=	𑌓	kā + 𑌟 -ṃ (11101)
𑌠	kaḥ	=	𑌓	kā + 𑌠 -ḥ (11102)
𑌡	k	=	𑌓	kā + 𑌡 MAAYYAA

One of the interesting features of Chakma writing is that CANDRABINDU (*cānaphupudā*) can be used together with ANUSVARA (*ekaphudā*) and VISARGA (*dviphudā*):

𑌛𑌛	aḥṁ	=	𑌛	ā + 𑌛	ḥ + 𑌛	ṁ
𑌛𑌛	aṃṁ	=	𑌛	ā + 𑌛	ṃ + 𑌛	ṁ
𑌛𑌛	uṃṁ	=	𑌛	u + 𑌛	ṃ + 𑌛	ṁ
𑌛𑌛	muṁ	=	𑌛	mā + 𑌛	u + 𑌛	ṁ

Glyph rendering of CANDRABINDU with other vowel signs (like 𑌔 *a* or 𑌕 *i*) is as yet unattested; nevertheless the encoding is clear.

5. Consonants with killed vowels and conjunct consonants. Like other Brahmic scripts, Chakma makes use of the MAAYYAA (*killer*) to invoke conjoined consonants. In the past, practice was much more common than it is today. Like the Myanmar script, Chakma is encoded with two vowel-killing characters in order to conform to modern user expectations. As shown above, most letters have their vowels killed with the use of the explicit MAAYYAA character:

𑌛	k	=	𑌓	kā + 𑌛	MAAYYAA
---	---	---	---	--------	---------

In 2001 an orthographic reform was recommended in the book *Cānmā pattham pāt* which would limit the standard repertoire of conjuncts to those composed with the five letters 𑌛 *yā*, 𑌛 *rā*, 𑌛 *lā*, 𑌛 *wā*, and 𑌛 *nā*. The four here are the most widely-accepted repertoire of conjuncts.

ya: X + ◻ VIRAMA + ဝ yā:

က ရ ဝ ယ င - ပ ခ ဓ ဟ ဇ - ဇ ဇ ဉ ဟ ဖ
က ဝ ဉ ဝ ဇ - ပ ဗ ဗ ဟ ဝ - ဝ မ ဟ ဝ ဝ

ra: X + ◻ VIRAMA + ဝ rā:

က ဝ ဝ ယ င - ပ ခ ဓ ဟ ဇ - ဇ ဇ ဉ ဟ ဖ
က ဝ ဉ ဝ ဇ - ပ ဗ ဗ ဟ ဝ - ဝ မ ဟ ဝ ဝ

la: X + ◻ VIRAMA + ဝ lā:

က လ ဝ ယ င - ပ ခ ဓ ဟ ဇ - ဇ ဇ ဉ ဟ ဖ
က လ ဝ ဉ ဝ ဇ - ပ ဗ ဗ ဟ ဝ - ဝ မ ဟ ဝ ဝ

wa: X + ◻ VIRAMA + ဝ wā:

က ဝ ဝ ယ င - ပ ခ ဓ ဟ ဇ - ဇ ဇ ဉ ဟ ဖ
က ဝ ဉ ဝ ဇ - ပ ဗ ဗ ဟ ဝ - ဝ မ ဟ ဝ ဝ

No separate conjunct forms of subjoined full-form *-yā* or *-rā* appear to exist. The fifth of these conjuncts, the *-na* conjunct, is exemplary of the orthographic shift which has taken place in Chakma.

na: X + ◻ VIRAMA + ဝ nā:

က ဝ ဝ ယ င - ပ ခ ဓ ဟ ဇ - ဇ ဇ ဉ ဟ ဖ
က ဝ ဉ ဝ ဇ - ပ ဗ ဗ ဟ ဝ - ဝ မ ဟ ဝ ဝ

While some writers would indeed write *kakna* (in ligating style) as ကက or (in subjoining style) as ကက, most now would probably expect it to be written as ကက. The ligating style of glyphs is now considered old-fashioned. Thus, taking the letter ဝ *mā* as the second element, while the glyph shapes ကက *kmā*, ကက *tmā*, ကက *nmā*, ကက *bbā*, ကက *mmā*, ကက *llā*, ကက *smā*, and ကက *hmā* are attested, most users now prefer the glyph shapes ကက *kmā*, ကက *tmā*, ကက *nmā*, ကက *bbā*, ကက *mmā*, ကက *llā*, ကက *smā*, and ကက *hmā*. Again, this distinction is stylistic and not orthographic.

As with Myanmar and Meetei Mayek, encoding a visible killer for modern users alongside an explicit conjoin-former permits the user to make specific choices about spelling more easily. Both the Myanmar encoding model and the Devanagari encoding model have been explained to the user community and feedback is that the Myanmar model fits the script better. (This is little surprise considering the close relationship between the Myanmar and Chakma scripts.)

In principle, nothing prevents the visible killer from appearing together with a stack. Thus while က *kā* + ◻ VIRAMA + ဝ *sā* gives က *ksā*, the sequence က *kā* + ◻ VIRAMA + ဝ *sā* + ◻ MAAYYAA could give က *ks*. Neither the string ◻ VIRAMA + ◻ MAAYYAA nor the string ◻ MAAYYAA + ◻ VIRAMA following a consonant would be meaningful, however, as both kill the inherent vowel, and “double-killing” a vowel makes no sense; either sequence should generate an error presentation.

The 2004 book *Phadagan* shows examples of the five conjuncts above together alongside conjuncts formed with ဝ *bā*, ဝ *mā*, and ဝ *hā*. These are all formed by simple subjoining.

ba: X + VIRAMA + ூ bā:

ூ ூ ூ ூ ூ ூ - ூ ூ ூ ூ ூ ூ - ூ ூ ூ ூ ூ ூ
 ூ ூ ூ ூ ூ ூ - ூ ூ ூ ூ ூ ூ - ூ ூ ூ ூ ூ ூ

ma: X + VIRAMA + ௃ mā:

௃ ௃ ௃ ௃ ௃ ௃ - ௃ ௃ ௃ ௃ ௃ ௃ - ௃ ௃ ௃ ௃ ௃ ௃
 ௃ ௃ ௃ ௃ ௃ ௃ - ௃ ௃ ௃ ௃ ௃ ௃ - ௃ ௃ ௃ ௃ ௃ ௃

ha: X + VIRAMA + ெ hā:

ெ ெ ெ ெ ெ ெ - ெ ெ ெ ெ ெ ெ - ெ ெ ெ ெ ெ ெ
 ெ ெ ெ ெ ெ ெ - ெ ெ ெ ெ ெ ெ - ெ ெ ெ ெ ெ ெ

In the 1982 book *Cāṅmār āg pudhi* a much wider range of conjunct pairs is shown, some of them with fairly complicated glyphs.

- ூ kkā = ூ kā + VIRAMA + ூ kā
- ௃ kṭā = ூ kā + VIRAMA + ௃ ṭā
- ௄ ktā = ூ kā + VIRAMA + ௄ tā
- ௅ kmā = ூ kā + VIRAMA + ௅ mā
- ெ kcā = ூ kā + VIRAMA + ெ cā (*conjunct shows old-style glyph*)
- ே ṅkā = ே ṅā + VIRAMA + ூ kā
- ை ṅgā = ே ṅā + VIRAMA + ை gā
- ௉ ccā = ெ cā + VIRAMA + ெ cā (*conjunct shows old-style glyph*)
- ொ cchā = ெ cā + VIRAMA + ொ chā (*conjunct shows old-style glyph*)
- ோ ṅcā = ௉ ṅā + VIRAMA + ெ cā (*conjunct shows old-style glyph*)
- ௌ ṅjā = ௉ ṅā + VIRAMA + ௌ jā
- ் ṅjhā = ௉ ṅā + VIRAMA + ் jhā
- ௎ ṭṭā = ௃ ṭā + VIRAMA + ௎ ṭā
- ௏ ttā = ௄ tā + VIRAMA + ௏ tā
- ௐ tmā = ௄ tā + VIRAMA + ௐ mā
- ௑ tthā = ௄ tā + VIRAMA + ௑ thā
- ௒ ddā = ௅ dā + VIRAMA + ௒ dā
- ௓ ddhā = ௅ dā + VIRAMA + ௓ dhā

𑄎𑄏	ntā	=	𑄎	nā	+	𑄏	VIRAMA	+	𑄏	tā
𑄎𑄏𑄏	nthā	=	𑄎	nā	+	𑄏	VIRAMA	+	𑄏𑄏	thā
𑄎𑄏𑄏	nmā	=	𑄎	nā	+	𑄏	VIRAMA	+	𑄏𑄏	mā
𑄏𑄏	ppā	=	𑄏	pā	+	𑄏	VIRAMA	+	𑄏	pā
𑄏𑄏	bbā	=	𑄏	bā	+	𑄏	VIRAMA	+	𑄏	bā
𑄏𑄏	mmā	=	𑄏	mā	+	𑄏	VIRAMA	+	𑄏	mā
𑄏𑄏	jjā	=	𑄏	jā	+	𑄏	VIRAMA	+	𑄏	jā
𑄏𑄏	lkā	=	𑄏	lā	+	𑄏	VIRAMA	+	𑄏	kā
𑄏𑄏	lgā	=	𑄏	lā	+	𑄏	VIRAMA	+	𑄏	gā
𑄏𑄏	llā	=	𑄏	lā	+	𑄏	VIRAMA	+	𑄏	lā
𑄏𑄏	ltā	=	𑄏	lā	+	𑄏	VIRAMA	+	𑄏	tā
𑄏𑄏	lpā	=	𑄏	lā	+	𑄏	VIRAMA	+	𑄏	pā
𑄏𑄏	schā	=	𑄏	sā	+	𑄏	VIRAMA	+	𑄏	chā (<i>conjunct shows old-style glyph</i>)
𑄏𑄏	stā	=	𑄏	sā	+	𑄏	VIRAMA	+	𑄏	tā
𑄏𑄏	skā	=	𑄏	sā	+	𑄏	VIRAMA	+	𑄏	kā
𑄏𑄏	spā	=	𑄏	sā	+	𑄏	VIRAMA	+	𑄏	pā
𑄏𑄏	smā	=	𑄏	sā	+	𑄏	VIRAMA	+	𑄏	mā
𑄏𑄏	hmā	=	𑄏	hā	+	𑄏	VIRAMA	+	𑄏	mā

5.1 Specific recommendation for Chakma fonts. In Chakma, the encoding model supports conjunct behaviour and Chakma fonts *by default* should display the subjoined form of letters when following virama, to ensure legibility. Whether a conjunct is required or not is part of the spelling of a word; it is not a stylistic issue. (We have seen no examples of conjuncts with more than one consonant, and while the encoding handles (in principle) any length of stacking examples of such would probably be spelling errors.)

6. Collating order. As an Indo-European language, the standard Brahmic sorting order applies to Chakma.

7. Character names. Consonant letter names use the typical Brahmic transliteration used in the UCS. Chakma letters have a descriptive name followed by a traditional Brahmic consonant. These latter are given in annotations to the character names.

8. Punctuation and digits. Alongside a | DANDA and || DOUBLE DANDA punctuation, Chakma has a unique 𑄏 QUESTION MARK, and a 𑄏 SECTION MARK. There is some variation in the glyphs for the SECTION MARK, some looking like flowers or leaves. A set of digits exists and is encoded, although Bengali digits are also used. The Tanchangya use Myanmar digits.

9. Linebreaking. Letters and digits behave as in Bengali. Both CHAKMA DANDA and CHAKMA DOUBLE DANDA behave as in Devanagari. The CHAKMA QUESTION MARK behaves like U+003F QUESTION MARK. The CHAKMA SECTION MARK behaves like U+2055 FLOWER PUNCTUATION MARK.

10. Unicode Character Properties.

```

11100;CHAKMA SIGN CANDRABINDU;Mn;0;NSM;;;;;N;;;;;
11101;CHAKMA SIGN ANUSVARA;Mn;0;NSM;;;;;N;;;;;
11102;CHAKMA SIGN VISARGA;Mc;0;L;;;;;N;;;;;
11103;CHAKMA LETTER AA;Lo;0;L;;;;;N;;;;;
11104;CHAKMA LETTER I;Lo;0;L;;;;;N;;;;;
11105;CHAKMA LETTER U;Lo;0;L;;;;;N;;;;;
11106;CHAKMA LETTER E;Lo;0;L;;;;;N;;;;;
11107;CHAKMA LETTER KAA;Lo;0;L;;;;;N;;;;;
11108;CHAKMA LETTER KHAA;Lo;0;L;;;;;N;;;;;
11109;CHAKMA LETTER GAA;Lo;0;L;;;;;N;;;;;
1110A;CHAKMA LETTER GHAA;Lo;0;L;;;;;N;;;;;
1110B;CHAKMA LETTER NGAA;Lo;0;L;;;;;N;;;;;
1110C;CHAKMA LETTER CAA;Lo;0;L;;;;;N;;;;;
1110D;CHAKMA LETTER CHAA;Lo;0;L;;;;;N;;;;;
1110E;CHAKMA LETTER JAA;Lo;0;L;;;;;N;;;;;
1110F;CHAKMA LETTER JHAA;Lo;0;L;;;;;N;;;;;
11110;CHAKMA LETTER NYAA;Lo;0;L;;;;;N;;;;;
11111;CHAKMA LETTER TTAA;Lo;0;L;;;;;N;;;;;
11112;CHAKMA LETTER TTHAA;Lo;0;L;;;;;N;;;;;
11113;CHAKMA LETTER DDAA;Lo;0;L;;;;;N;;;;;
11114;CHAKMA LETTER DDHAA;Lo;0;L;;;;;N;;;;;
11115;CHAKMA LETTER NNAA;Lo;0;L;;;;;N;;;;;
11116;CHAKMA LETTER TAA;Lo;0;L;;;;;N;;;;;
11117;CHAKMA LETTER THAA;Lo;0;L;;;;;N;;;;;
11118;CHAKMA LETTER DAA;Lo;0;L;;;;;N;;;;;
11119;CHAKMA LETTER DHAA;Lo;0;L;;;;;N;;;;;
1111A;CHAKMA LETTER NAA;Lo;0;L;;;;;N;;;;;
1111B;CHAKMA LETTER PAA;Lo;0;L;;;;;N;;;;;
1111C;CHAKMA LETTER PHAA;Lo;0;L;;;;;N;;;;;
1111D;CHAKMA LETTER BAA;Lo;0;L;;;;;N;;;;;
1111E;CHAKMA LETTER BHAA;Lo;0;L;;;;;N;;;;;
1111F;CHAKMA LETTER MAA;Lo;0;L;;;;;N;;;;;
11120;CHAKMA LETTER YYAA;Lo;0;L;;;;;N;;;;;
11121;CHAKMA LETTER YAA;Lo;0;L;;;;;N;;;;;
11122;CHAKMA LETTER RAA;Lo;0;L;;;;;N;;;;;
11123;CHAKMA LETTER LAA;Lo;0;L;;;;;N;;;;;
11124;CHAKMA LETTER WAA;Lo;0;L;;;;;N;;;;;
11125;CHAKMA LETTER SAA;Lo;0;L;;;;;N;;;;;
11126;CHAKMA LETTER HAA;Lo;0;L;;;;;N;;;;;
11127;CHAKMA VOWEL SIGN A;Mn;0;NSM;;;;;N;;;;;
11128;CHAKMA VOWEL SIGN I;Mn;0;NSM;;;;;N;;;;;
11129;CHAKMA VOWEL SIGN II;Mn;0;NSM;;;;;N;;;;;
1112A;CHAKMA VOWEL SIGN U;Mn;0;NSM;;;;;N;;;;;
1112B;CHAKMA VOWEL SIGN UU;Mn;0;NSM;;;;;N;;;;;
1112C;CHAKMA VOWEL SIGN E;Mc;0;L;;;;;N;;;;;
1112D;CHAKMA VOWEL SIGN AI;Mn;0;NSM;;;;;N;;;;;
1112E;CHAKMA VOWEL SIGN O;Mn;0;NSM;11131 11127;;;;;N;;;;;
1112F;CHAKMA VOWEL SIGN AU;Mn;0;NSM;11132 11127;;;;;N;;;;;
11130;CHAKMA VOWEL SIGN OI;Mn;0;NSM;;;;;N;;;;;
11131;CHAKMA O MARK;Mn;0;NSM;;;;;N;;;;;
11132;CHAKMA AU MARK;Mn;0;NSM;;;;;N;;;;;
11133;CHAKMA VIRAMA;Mn;9;NSM;;;;;N;;;;;
11134;CHAKMA MAAYYAA;Mn;9;NSM;;;;;N;;;;;
11136;CHAKMA DIGIT ZERO;Nd;0;L;0;0;0;N;;;;;
11137;CHAKMA DIGIT ONE;Nd;0;L;1;1;1;N;;;;;
11138;CHAKMA DIGIT TWO;Nd;0;L;2;2;2;N;;;;;
11139;CHAKMA DIGIT THREE;Nd;0;L;3;3;3;N;;;;;
1113A;CHAKMA DIGIT FOUR;Nd;0;L;4;4;4;N;;;;;
1113B;CHAKMA DIGIT FIVE;Nd;0;L;5;5;5;N;;;;;
1113C;CHAKMA DIGIT SIX;Nd;0;L;6;6;6;N;;;;;
1113D;CHAKMA DIGIT SEVEN;Nd;0;L;7;7;7;N;;;;;
1113E;CHAKMA DIGIT EIGHT;Nd;0;L;8;8;8;N;;;;;
1113F;CHAKMA DIGIT NINE;Nd;0;L;9;9;9;N;;;;;
11140;CHAKMA SECTION MARK;Po;0;L;;;;;N;;;;;
11141;CHAKMA DANDA;Po;0;L;;;;;N;;;;;
11142;CHAKMA DOUBLE DANDA;Po;0;L;;;;;N;;;;;
11143;CHAKMA QUESTION MARK;Po;0;L;;;;;N;;;;;

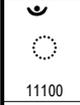
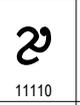
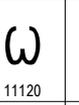
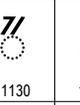
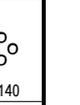
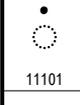
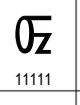
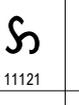
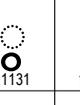
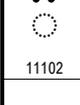
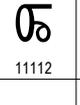
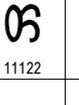
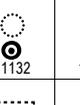
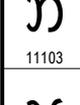
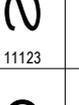
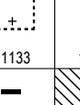
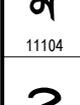
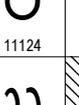
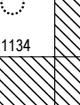
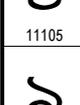
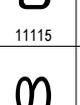
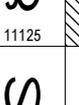
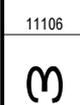
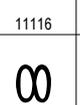
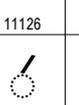
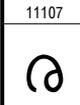
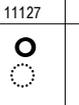
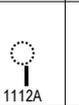
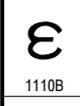
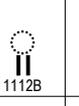
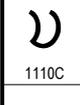
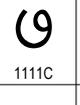
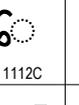
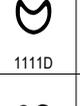
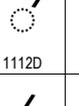
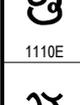
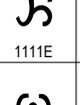
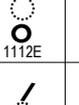
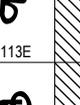
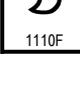
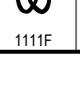
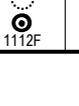
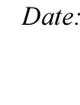
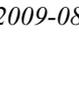
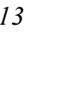
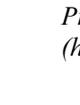
```

11. Bibliography.

Cānmā, Cirajyoti and Maṅgal Cāngmā. 1982. *Cānmār āg pudhi (Chakma primer)*. Rānamāṭi: Cānmābhāṣā Prakāśanā Pariṣad.

Khisa, Bhagadatta. 2001. *Cāṅmā pattham pāt = Chakma primer*. Rāṅmāṅṅi: Tribal Cultural Institute (TCI).
Singā. 2004. *Phagadāṅ*.

12. Acknowledgements. This project was made possible in part by a grant from the U.S. National Endowment for the Humanities, which funded the Universal Scripts Project (part of the Script Encoding Initiative at UC Berkeley) in respect of the Chakma encoding. Any views, findings, conclusions or recommendations expressed in this publication do not necessarily reflect those of the National Endowment of the Humanities.

	1110	1111	1112	1113	1114
0	 11100	 11110	 11120	 11130	 11140
1	 11101	 11111	 11121	 11131	 11141
2	 11102	 11112	 11122	 11132	 11142
3	 11103	 11113	 11123	 11133	 11143
4	 11104	 11114	 11124	 11134	
5	 11105	 11115	 11125		
6	 11106	 11116	 11126	 11136	
7	 11107	 11117	 11127	 11137	
8	 11108	 11118	 11128	 11138	
9	 11109	 11119	 11129	 11139	
A	 1110A	 1111A	 1112A	 1113A	
B	 1110B	 1111B	 1112B	 1113B	
C	 1110C	 1111C	 1112C	 1113C	
D	 1110D	 1111D	 1112D	 1113D	
E	 1110E	 1111E	 1112E	 1113E	
F	 1110F	 1111F	 1112F	 1113F	

Various signs

- 11100 ◌ CHAKMA SIGN CANDRABINDU
= caanaphupudaa
- 11101 ◌ CHAKMA SIGN ANUSVARA
= ekaphudaa
- 11102 ◌ CHAKMA SIGN VISARGA
= dviphudaa

Independent vowels

- 11103 𑄢 CHAKMA LETTER AA
= pichapujhaa aa
- 11104 𑄣 CHAKMA LETTER I
= delabhaangagaa i
- 11105 𑄤 CHAKMA LETTER U
= bacacu u
- 11106 𑄥 CHAKMA LETTER E
= lejaubaa e

Consonants

- 11107 𑄦 CHAKMA LETTER KAA
= cucyaangyaa kaa
- 11108 𑄧 CHAKMA LETTER KHAA
= grajaangyaa khaa
- 11109 𑄨 CHAKMA LETTER GAA
= caandyyaa gaa
- 1110A 𑄩 CHAKMA LETTER GHAA
= tinaddaalyaa ghaa
- 1110B 𑄪 CHAKMA LETTER NGAA
= cilaama ngaa
- 1110C 𑄫 CHAKMA LETTER CAA
= dvibhalyaa caa
- 1110D 𑄬 CHAKMA LETTER CHAA
= majaraa chaa
- 1110E 𑄭 CHAKMA LETTER JAA
= dvipadala haa
- 1110F 𑄮 CHAKMA LETTER JHAA
= uraauraa jhaa
- 11110 𑄯 CHAKMA LETTER NYAA
= silaacyaa nyaa
- 11111 𑄰 CHAKMA LETTER TTA
= dviyaadaat ttaa
- 11112 𑄱 CHAKMA LETTER TTHAA
= phudaadviaat thaa
- 11113 𑄲 CHAKMA LETTER DDAA
= aadudaangaat ddaa
- 11114 𑄳 CHAKMA LETTER DDHAA
= lejabharaat ddhaa
- 11115 𑄴 CHAKMA LETTER NNAA
= petttuyaa nnaa
- 11116 𑄵 CHAKMA LETTER TAA
= ghangadaat taa
- 11117 𑄶 CHAKMA LETTER THAA
= jagadaat thaa
- 11118 𑄷 CHAKMA LETTER DAA
= dolaniit daa
- 11119 𑄸 CHAKMA LETTER DHAA
= talamuyaat dhaa
- 1111A 𑄹 CHAKMA LETTER NAA
= phaarabaanyaa naa
- 1111B 𑄺 CHAKMA LETTER PAA
= paalyaa paa
- 1111C 𑄻 CHAKMA LETTER PHAA
= ubaraphudaa phaa

- 1111D 𑄼 CHAKMA LETTER BAA
= ubaramuyaa baa
- 1111E 𑄽 CHAKMA LETTER BHAA
= ciraddaalyaa bhaa
- 1111F 𑄾 CHAKMA LETTER MAA
= bugatpadalaa maa
- 11120 𑄿 CHAKMA LETTER YYAA
= cimayyaa yyaa
- 11121 𑅀 CHAKMA LETTER YAA
= jilyaa yaa
- 11122 𑅁 CHAKMA LETTER RAA
= dvidaayyaa raa
- 11123 𑅂 CHAKMA LETTER LAA
= talamuyaa laa
- 11124 𑅃 CHAKMA LETTER WAA
= bajhonyaa waa
- 11125 𑅄 CHAKMA LETTER SAA
= bhudibukyaa saa
- 11126 𑅅 CHAKMA LETTER HAA
= ubaramuyaa haa

Dependent vowel signs

- 11127 𑅆 CHAKMA VOWEL SIGN A
= ubaratulyaa a
- 11128 𑅇 CHAKMA VOWEL SIGN I
= bahryaa i
- 11129 𑅈 CHAKMA VOWEL SIGN II
= baaniiphadaa ii
- 1112A 𑅉 CHAKMA VOWEL SIGN U
= ekattaana u
- 1112B 𑅊 CHAKMA VOWEL SIGN UU
= dviittaana uu
- 1112C 𑅋 CHAKMA VOWEL SIGN E
= ekaara e
- 1112D 𑅌 CHAKMA VOWEL SIGN AI
= delabhaanga ai
- 1112E 𑅍 CHAKMA VOWEL SIGN O
= okaara o
≡ 11131 ◌ 11127 𑅆
- 1112F 𑅎 CHAKMA VOWEL SIGN AU
= aukaara au
≡ 11132 ◌ 11127 𑅆
- 11130 𑅏 CHAKMA VOWEL SIGN OI
= oikaara oi
- 11131 ◌ CHAKMA O MARK
- 11132 ◌ CHAKMA AU MARK

Various signs

- 11133 ◻ CHAKMA VIRAMA
• used to form conjuncts
→ 1039 ◻ myanmar sign virama
- 11134 ◌ CHAKMA MAAYYAA
• killer
→ 103A ◌ myanmar sign asat

Digits

- 11136 0 CHAKMA DIGIT ZERO
- 11137 𑅐 CHAKMA DIGIT ONE
- 11138 𑅑 CHAKMA DIGIT TWO
- 11139 𑅒 CHAKMA DIGIT THREE
- 1113A 𑅓 CHAKMA DIGIT FOUR
- 1113B 𑅔 CHAKMA DIGIT FIVE

- 1113C 𑄎 CHAKMA DIGIT SIX
1113D 𑄏 CHAKMA DIGIT SEVEN
1113E 𑄐 CHAKMA DIGIT EIGHT
1113F 𑄑 CHAKMA DIGIT NINE

Punctuation

- 11140 𑄒 CHAKMA SECTION MARK
= phulacihna
11141 𑄓 CHAKMA DANDA
= ekacilyaa
11142 𑄔 CHAKMA DOUBLE DANDA
= dvicilyaa
11143 𑄕 CHAKMA QUESTION MARK
= pujhaar

DR. G. A. Grierson সংগৃহীত চাকমা বর্ণ (১৯০৩ খ্র:) এবং অন্তর্লিঙ্গ জন্মবর্ষে তার মতামত

The following account of the Chākma alphabet is based on information provided by Dewan Kristo Chandra, a gentleman of Chākma nationality, and forwarded to me by Mr. J. A. Cave-Browne, Assistant Commissioner, Chittagong Hill Tracts.

The Chākma alphabet is as follows:—

က	က	ဂ	ဂ	ဣ
ka	khā	gā	ghā	nā
ဣ	ဣ	ဣ	ဣ	ဣ
chā (sā)	chhā	jā	jhā	nā
ဣ	ဣ	ဣ	ဣ	ဣ
ta	thā	dā	dhā	nā
က	က	ဂ	ဂ	ဣ
lā	lhā	dā	dhā	nā
ဣ	ဣ	ဣ	ဣ	ဣ
pā	phā	bā	bhā	mā
ဣ	ဣ	ဣ	ဣ	ဣ
yā	rā	lā	ra	shā
က	က	ဂ		
kā	khā	d.		

Bengali. 322 BENGALI.

The most important point to notice in this alphabet is that the vowel inherent in each consonant is, not *a* as in other Indian languages, but *ā*. Note also that the initial form (there is, of course, no non-initial form) of *d* is treated as a consonant, much as the letter *alif* is treated as a consonant in Arabic.

For purposes of comparison, I here give the usual Burmese forms of the consonants:—

က	ka,	ခ	khā,	ဂ	gā,	ဃ	ghā,	ဣ	nā,
ဣ	chā,	ဣ	chhā,	ဣ	jā,	ဣ	jhā,	ဣ	nā,
ဣ	ta,	ဣ	thā,	ဣ	dā,	ဣ	dhā,	ဣ	na,
က	ta,	က	thā,	ဂ	dā,	ဣ	dhā,	ဣ	na,
ဣ	pā,	ဣ	phā,	ဣ	bā,	ဣ	bhā,	ဣ	mā,
ဣ	ya,	ဣ	ra,	ဣ	lā,	ဣ	ra		
ဣ	ka,	ဣ	khā						

Figure 1. Chakma chart from Grierson's *Linguistic Survey of India*, 1903.

ৰাঙ্গামাটিস্থ উপজাতীয় সাংস্কৃতিক ইনষ্টিটিউট ২০০১ সালে "চাঙমা পত্ৰম পাত" গ্ৰন্থে প্ৰকাশিত

চাক্ৰমা বৰ্ণমালা				
ᱠ চুচাঙা - ক	ᱡ চুচাঙা - খ	ᱢ চান্দা - গ	ᱣ তিনখালা - ঘ	ᱤ চিলাম - ঙ
ᱥ বিখালা - চ	ᱦ মোছাঙা - ছ	ᱧ বিপদলা - জ	ᱨ উৰাউবি - ঝ	ᱩ চিলোচা - ঞ
ᱪ বিহাদা - ট	ᱫ ফুদাদিয়া - ঠ	ᱬ হাদুভাঙা - ড	ᱭ শেঙ্গুভাঙা - ঢ	ᱮ পেঙ্গুয়া - গ
ᱯ ঘৰুদা - তা	ᱰ হুগদা - থ	ᱱ দোলনিয়া - দা	ᱲ তলমুয়া - ধ	ᱳ ফাৰবাণ্ডা - না
ᱴ পালা - পা	ᱵ উবরপদা - ফ	ᱶ উবরমুয়া - বা	ᱷ চেইৰখাশা - জ	ᱸ হুগতপদা - মা
ᱹ জিলা - যা	ᱺ বিদাঙা - ঝ	ᱻ তলমুয়া - না	ᱼ বাখোন্দা - ওয়া	ᱽ উবরমুয়া - হ
᱾ জুদিবুকা - সা	᱿ চিমোছা - ষা			
ᱠ অ	ᱡ আ	ᱢ ই	ᱣ উ	ᱤ এ
ᱥ ও	ᱦ ঔ			

৬৬

চিহ্নৰ ব্যৱহাৰ					
বৰ্ণ		চিহ্ন			
ᱠ	উবর তুলো	ˊ	ᱠ	ক	যেমন ᱠᱢᱣᱟ
ᱠ	মাঙা	ˉ	ᱠ	ক্	যেমন ᱠᱢᱣᱟ
ᱠ	বানো	ᱚ	ᱠ	কি	যেমন ᱠᱢᱣᱟ
ᱠ	একটন দিলে	ᱚ	ᱠ	ক্	যেমন ᱠᱢᱣᱟ
ᱠ	এ-কাৰ দিলে	ᱚ	ᱠ	কে	যেমন ᱠᱢᱣᱟ
ᱠ	ও-কাৰ দিলে	ᱚ	ᱠ	কো	যেমন ᱠᱢᱣᱟ
ᱠ	দেল ডাঙিলে	ᱚ	ᱠ	কাই	যেমন ᱠᱢᱣᱟ
ᱠ	য়া - দিলে	ᱚ	ᱠ	ক্যা	যেমন ᱠᱢᱣᱟ
ᱠ	ৰা - দিলে	ᱚ	ᱠ	ক্ৰা	যেমন ᱠᱢᱣᱟ
ᱠ	লা - দিলে	ᱚ	ᱠ	ক্ৰা	যেমন ᱠᱢᱣᱟ
ᱠ	ওয়া - দিলে	ᱚ	ᱠ	কোয়া	যেমন ᱠᱢᱣᱟ
ᱠ	হা - দিলে	ᱚ	ᱠ	কাহ	যেমন ᱠᱢᱣᱟ
ᱠ	না - দিলে	ᱚ	ᱠ	ক্ৰা	যেমন ᱠᱢᱣᱟ
ᱠ	একফুদা দিলে	˙	ᱠ	কাং	যেমন ᱠᱢᱣᱟ
ᱠ	বিফুদা দিলে	˙˙	ᱠ	কাঃ	যেমন ᱠᱢᱣᱟ
ᱠ	চান ফুদা দিলে	˙˙˙	ᱠ	কাঁ	যেমন ᱠᱢᱣᱟ

৬৭

Figure 2. Charts taken from a paper written by Mr Sugata Chakma of the Tribal Cultural Institute, on "the Primary classification of languages".

প্ৰিয়তম্যে নৈব

(৩)

গেহে তুগেই হেঁ হেঁ
 নত্বে হেঁ হু খিল হেঁ
 তখত নীত্বে হেঁ হেঁ
 তীদেহে হুইত্বে হেঁ

(৪)

গল হেহে গল তুগেই
 নুগেই হেহে নুগেই
 হেঁ হেঁ হেঁ হেঁ হেঁ
 হেঁ হেঁ হেঁ হেঁ হেঁ

(৫)

খিল হেহে তুগেই হেঁ
 হেঁ, নীত্বে হেঁ হেঁ
 হেঁ-হু হেঁ, নুগেই হেঁ
 হেঁ হেঁ হেঁ হেঁ, হেঁ হেঁ হেঁ হেঁ হেঁ
 হেঁ-হু হেঁ, হেঁ

Figure 3. Example of poetry from the book *Phagadān*, 2004.

চাক্ৰমা বৰ্ণমালা				
চূচ্যাঙা - কা	গুজ্যাঙা - খা	চান্দ্যা - গা	তিনধালা - ঘা	চিলাম - ঙা
দ্বিধালা - চা	মোজৰ্জ্যা - ছা	দ্বিপদলা - জা	উরাউরি - বা	চিলোচ্যা - ঞা
দ্বিহাদাত্ - টা	ফুদাদিয়াত্ - ঠা	হাদুভাঙাত্ - ডা	লেজভৰাত - ঢা	পেতুয়া - ণা
ঘঙদাত্ - তা	জগদাত্ - থা	দোলনীয়ত - দা	তলমুয়াত্ - ধা	ফাৰবান্যা - না
পাল্যা - পা	উবৰপদলা - ফা	উবৰমুয়া - বা	চেইৰধালা - ভা	বুগতপদলা - মা
জিল্যা - যা	দ্বিদাৰ্জ্যা - রা	তলমুয়া - লা	বাবোন্যা - ওয়া	উবৰমুয়া - হা
ভুদিবুক্যা - সা	চিমোজ্যা - য়া			
অ	আ	ই	উ	এ
ও	ঐ			

Figure 4. Alphabet chart from Khisa 2001.

চিহ্নের ব্যবহার

বর্ণ		চিহ্ন				
৓	উবর তুল্যে	্	৓	ক	যেমন	ক্‌শ্‌খ্‌শ্‌
৓	মাজ্য	—	৓̄	ক্	যেমন	ক্‌খ্‌ক্‌
৓	বান্যে	০	৓̆	কি	যেমন	কি̆খ্‌কি̆
৓	একটান দিলে	।	৓।	কু	যেমন	কু।খ্‌কু।
৓	এ-কার দিলে	ে	৓ে	কে	যেমন	কেেখ্‌কেে
৓	ও-কার দিলে	ে	৓ে	কো	যেমন	কোেখ্‌কোে
৓	দেল ভাঙিলে	্	৓্	কাই	যেমন	কাই্‌
৓	য়া - দিলে	্	৓্	ক্যা	যেমন	ক্যা্‌
৓	রা - দিলে	্	৓্	ক্রা	যেমন	ক্রা্‌
৓	লা - দিলে	্	৓্	ক্রা	যেমন	ক্রা্‌
৓	ওয়া - দিলে	্	৓্	কোয়া	যেমন	কোয়া্‌
৓	হা - দিলে	্	৓্	কাহ্	যেমন	কাহ্‌
৓	না - দিলে	্	৓্	ক্রা	যেমন	ক্রা্‌
৓	একফুদা দিলে	·	৓̇	কাং	যেমন	কাং̇
৓	দ্বিফুদা দিলে	·	৓̈	কাঃ	যেমন	কাঃ̈
৓	চান ফুদা দিলে	্	৓̈	কাঁ	যেমন	কাঁ̈

- ১। চাকমা বর্ণমালায় কোন যুক্তাক্ষর থাকবে না। তবে য, র, ল, ব, হ ও ন ফলা থাকবে। যেমনঃ
ঐ̈, ঐ̈, ঐ̈, ঐ̈ ইত্যাদি।
- ২। দীর্ঘ ই-কার, দীর্ঘ উ-কার থাকবে না। উ (৫) বর্ণের পরিবর্তে ঐ - উ, ই (২) বর্ণের পরিবর্তে ঐ - ই এবং এ (৯) বর্ণের পরিবর্তে ঐ - এ লিখতে হবে।
 হ্রা (০৫) বর্ণটির ব্যবহার হবে না।
- ৩। অংকের সংখ্যাগুলি বাংলা সংখ্যার ন্যায় লিখতে হবে- ১, ২, ৩, ৪, ৫, ৬, ৭, ৮, ৯, ১০

Figure 5. Chart of vowel signs and conjuncts from Khisa 2001.

চাওমার আগ্ পুথি
 ঠা ঠা ঠা ঠা - (যুক্ত অক্ষর)
 (জদা অক্ষর)

৪৭

চাওমা বর্ণমালার যুক্ত অক্ষরের আকার নিম্নরূপ
 অবশ্য এগুলি ছাড়াও চাওমা বর্ণমালায় আরোও অনেক
 যুক্ত অক্ষরের ব্যবহার আছে।

^ ক	^ খ	^ ক	^ ক	^ ক	^ ক
^ গ	^ ঘ	^ গ	^ গ	^ গ	^ গ
^ গ	^ ঘ	^ গ	^ গ	^ গ	^ গ
^ গ	^ ঘ	^ গ	^ গ	^ গ	^ গ
^ গ	^ ঘ	^ গ	^ গ	^ গ	^ গ
^ গ	^ ঘ	^ গ	^ গ	^ গ	^ গ

Figure 6. Chart with old-style conjuncts from Caimā 1982.

—	,			‡	=
^ জড়া	^ জিন্নান্	^ একচিল্যা	^ দ্বিচিল্যা	^ পুঝার্	^ সং
(-)	((-))	<->	//-//		
^ একবান	^ দ্বিবান	^ তিনবান	^ উদ্ধার		
⊙	+	-	×	÷	
^ ফুলচিহ্ন	^ এগত্তর	^ ফারক	^ দুনা	^ ভাগ	
					
০৯ ৩৯৬৬ - গননা হরফ।					
৪	২	৬	৭	৮	
^ এক	^ দুই	^ তিন	^ চের	^ পাচ	
৪	৭	৮	০৪	০	
^ ছ	^ সাত	^ আট	^ ন	^ শূন্য	

Figure 7. Chart with punctuation and digits from Cānmā 1982. Shown are: — EM DASH, , COMMA, | CHAKMA DANDA, || CHAKMA DOUBLE DANDA, ‡ CHAKMA QUESTION MARK, = EQUALS SIGN; (LEFT PARENTHESIS,) RIGHT PARENTHESIS, ((LEFT DOUBLE PARENTHESIS,)) RIGHT DOUBLE PARENTHESIS, < LESS-THAN SIGN, > GREATER-THAN SIGN (unless these are single guillemets < > or angle brackets < >), SOLIDUS repeated (the name given is *udkār* 'quotation' so these could likely be quotation marks ""); ⊙ CHAKMA SECTION MARK, + PLUS SIGN, - MINUS SIGN, × MULTIPLICATION SIGN, and ÷ DIVISION SIGN.

ཀ k	ཁ kh	ག g	མ gh	ང n
པ c	ཅ ch	ཉ j	ཞ jh	ཎ n
ཏ t	ཐ th	ཛ d	ཌ dh	ཎ n
ཏ t	ཐ th	ཛ d	ཌ dh	ཎ n
ཕ p	ཆ ph	བ b	མ bh	མ m
ཉ y	ཏ r	ལ l	ས s	ཏ h
ཡ y				
འ a	ཨ i	ཨ u	ཨ é	ཨ o
ཀ kā	ཀ ka	ཀ ki	ཀ ku	ཀ ké
ཀ kai	ཀ kua	ཀ ko	ཀ kāñ	ཀ kā
ཀ kyā	ཀ krā	ཀ klā	ཀ knā	ཀ k
༡ 1	༢ 2	༣ 3	༤ 4	༥ 5
༦ 6	༧ 7	༨ 8	༩ 9	༠ 0

Figure 9. Chakma alphabet chart from Bernot 1972.

A. Administrative

1. Title

Proposal for encoding the Chakma script in the UCS

2. Requester's name

UC Berkeley Script Encoding Initiative (Universal Scripts Project)

3. Requester type (Member body/Liaison/Individual contribution)

Liaison contribution.

4. Submission date

2009-08-13

5. Requester's reference (if applicable)

6. Choose one of the following:

6a. This is a complete proposal

No.

6b. More information will be provided later

Yes.

B. Technical – General

1. Choose one of the following:

1a. This proposal is for a new script (set of characters)

Yes.

1b. Proposed name of script

Chakma.

1c. The proposal is for addition of character(s) to an existing block

No.

1d. Name of the existing block

2. Number of characters in proposal

67.

3. Proposed category (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; G-Obscure or questionable usage symbols)

Category A.

4a. Is a repertoire including character names provided?

Yes.

4b. If YES, are the names in accordance with the "character naming guidelines" in Annex L of P&P document?

Yes.

4c. Are the character shapes attached in a legible form suitable for review?

Yes.

5a. Who will provide the appropriate computerized font (ordered preference: True Type, or PostScript format) for publishing the standard?

Michael Everson and Hangendra Chakma.

5b. If available now, identify source(s) for the font (include address, e-mail, ftp-site, etc.) and indicate the tools used:

Michael Everson, Fontographer.

6a. Are references (to other character sets, dictionaries, descriptive texts etc.) provided?

Yes.

6b. Are published examples of use (such as samples from newspapers, magazines, or other sources) of proposed characters attached?

Yes.

7. 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)?

Yes.

8. 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/Public/UNIDATA/UnicodeCharacterDatabase.html> and associated Unicode Technical Reports for information needed for consideration by the Unicode Technical Committee for inclusion in the Unicode Standard.

See above.

C. Technical – Justification

1. Has this proposal for addition of character(s) been submitted before? If YES, explain.

No.

2a. 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.

2b. If YES, with whom?

Hagendra Chakma, Provungshu Chakma, John Clifton, Keisuke Huziwara, Pragya Joyoti, Saikat Khisa, Helen Leake, Chandra Roy

2c. 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?

People living in Bangladesh and in India.

4a. The context of use for the proposed characters (type of use; common or rare)

Common.

4b. Reference

5a. Are the proposed characters in current use by the user community?

Yes.

5b. If YES, where?

In Bangladesh and in India.

6a. After giving due considerations to the principles in the P&P document must the proposed characters be entirely in the BMP?

Yes.

6b. If YES, is a rationale provided?

Yes.

6c. If YES, reference

Contemporary use and accordance with the Roadmap.

7. Should the proposed characters be kept together in a contiguous range (rather than being scattered)?

Yes.

8a. Can any of the proposed characters be considered a presentation form of an existing character or character sequence?

No.

8b. If YES, is a rationale for its inclusion provided?

8c. If YES, reference

9a. Can any of the proposed characters be encoded using a composed character sequence of either existing characters or other proposed characters?

No.

9b. If YES, is a rationale for its inclusion provided?

9c. If YES, reference

10a. Can any of the proposed character(s) be considered to be similar (in appearance or function) to an existing character?

No.

10b. If YES, is a rationale for its inclusion provided?

10c. If YES, reference

11a. Does the proposal include use of combining characters and/or use of composite sequences (see clauses 4.12 and 4.14 in ISO/IEC 10646-1: 2000)?

No.

11b. If YES, is a rationale for such use provided?

11c. If YES, reference

11d. Is a list of composite sequences and their corresponding glyph images (graphic symbols) provided?

No.

11e. If YES, reference

12a. Does the proposal contain characters with any special properties such as control function or similar semantics?

No.

12b. If YES, describe in detail (include attachment if necessary)

13a. Does the proposal contain any Ideographic compatibility character(s)?

No.

13b. If YES, is the equivalent corresponding unified ideographic character(s) identified?