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

1. Introduction. The Mandaic script is used to write a dialect of Eastern Aramaic, which, in its classical form, is currently used as the liturgical language of the Mandaean religion. A living language descended from Classical Mandaic is spoken by a small number of people living in and around Ahvaz, Khūzestān, in southwestern Iran; speakers are also found in emigrant communities in Sweden, Australia, and the United States. There is a considerable amount of Iranian influence on the lexicon of Classical Mandaic, and Arabic and Persian influence on the grammar and lexicon of the contemporary dialect. The script itself is likely derived from the Parthian chancery script.

2. Structure. Mandaic is a right-to-left script. It is a true alphabet, using letters regularly for vowels rather than as the *matres lectionis* from which they derived. The three diacritical marks are used in teaching materials to differentiate vowel quality. At present, at least, the rule is that they may be omitted from ordinary text. In this regard they are very like the Arabic *fatha*, *kasra*, and *damma* or the Hebrew vowel points.

- The AFFRICATION MARK ◌ is used to extend the character set for foreign sounds (whether affrication, lenition, or another sound). See figures 8 and 9. Compare:

ـﻊ	G	*g*	ـﻊ	G̱	*γ*
ﺪ	D	*d*	ﺪ	Ḏ	*δ*
ـﺳ	H	*h*	ـﺳ	Ḥ	*ḥ*
ﻞ	Ṭ	*ṭ*	ﻞ	Ṭ	*ẓ*
ﻴ	K	*k*	ﻴ	Ḵ	*χ*
ﻮﻭ	P	*p*	ﻮﻭ	P̱	*f*
ﻉﻭ	Ṣ	*ṣ*	ﻉﻭ	Ṣ	*ž?*
ﻗﻭ	Š	*š*	ﻗﻭ	Š̱	*č, ǰ*
ﻞﻭ	T	*t*	ﻞﻭ	Ṯ	*θ*

- The VOCALIZATION MARK ◌ is used to distinguish vowel quality of *halqa*, *ušenna*, and *aksa* (Hebrew *alef*, *waw*, *yod*):

ﻜﺤ	BA	*bā*	ﻜﺤ	BA̱	*ba*
ﻜﺪ	BU	*bu*	ﻜﺤ	BU̱	*bo*
ـﻜﻌ	BI	*bi*	ـﻜﻌ	BI̱	*be*

- The GEMINATION MARK ◌ is used to indicate what native writers call a "hard" pronunciation:

ﺤﻴﻭ	AḴA	*ekka*	'there is'	ﺤﻴﻭﻗﻭ	ŠIṆA	*šenna*	'tooth',
ﺤﺘ	LIḄA	*lebba*	'heart'	ﺤﺤ	RḄH	*rabba*	'great'

1

3. Joining behaviour. Mandaic has fully-developed joining behaviour. The table below shows the joining forms.

Dual-joining Mandaic Characters

Character	X_n	X_r	X_m	X_l
AB	ڪ	ڪ	ڪ	ڪ
AG	ع	ع	ع	ع
AD	ڪ	ڪ	ڪ	ڪ
AH	ـس	ـس	ـس	ـس
USHENNA	د	د	د	د
IT	ڪ	ڪ	ڪ	ڪ
ATT	⊥	⊥	⊥	⊥
AK	ע	ע	ע	ע
AL	⅃	⅃	⅃	⅃
AM	ﺝ	ﺝ	ﺝ	ﺝ
AN	ν	ν	ν	ν
AS	ص	ص	ص	ص
AP	y	y	y	y
ASZ	ﺵ	ﺵ	ﺵ	ﺵ
AQ	ﺝ	ﺝ	ﺝ	ﺝ
AR	ﺣ	ﺣ	ﺣ	ﺣ
AT	ﺟ	ﺟ	ﺟ	ﺟ

Right-joining Mandaic Characters

Character	X_n	X_r
HALQA	ه	ﻪ
AZ	ǀ	ǂ
IN	ه	ﻪ
AKSA	�	ﺣ
ASH	٩ه	٩ه

Non-joining Mandaic Characters

Character	X_n
DUSHENNA	ﺣ
KAD	ﺣﻳ
AIN	ع

U+0856 ﺣ MANDAIC LETTER DUSHENNA (also called *adu*), transliterated *ḏ*, is derived from an old ligature of U+0843 ﺣ MANDAIC LETTER AD *d* and U+0849 ﺣ MANDAIC LETTER AKSA *i*, used in Aramaic to write the relative pronoun and the genitive exponent *di*. It is used as an undecomposable letter in its own right (like Danish *æ*), usually written proclitically as the first letter of a word, or written alone. Note that AD and AKSA ligate normally: ﺣ *di* = ﺣ *i* + ﺣ *d* ← (reading right to left).

Similarly, U+0857 ﺣﻳ MANDAIC LETTER KAD is derived from an old ligature of U+084A ﻳ MANDAIC LETTER AK *k* and U+0856 ﺣ MANDAIC LETTER DUSHENNA *di*, used in Aramaic to write the word *kḏ* 'when, as, like'; compare Hebrew כדי *kǝdi*. It is also used as an undecomposable letter in its own right, usually written alone. While AK is dual joining, KAD never joins with a preceding character; the joining behaviour is different between the two despite the origin of KAD in AK + DUSHENNA. Compare ﺣﻳ*pkḏ* = ﺣ *ḏ* + ﻳ *k* + y *p* ← (where AP and AK join as normal and AK and DUSHENNA do not join) with ﺣﻳy *pkḏ* = ﺣﻳ *kḏ* + y *p* ← (where AP and KAD do not join) . Note also the similarity of KAD and the syllable *ki*: ﺣﻳ *ki* = ﺣ *i* + ﻳ *k* ← (reading right to left).

U+0858 ‫ع‬ MANDAIC LETTER AIN is a borrowing from U+0639 ‫ع‬ ARABIC LETTER AIN and as noted above does not combine with other characters.

4. Punctuation. Sentence punctuation is used, rather sparsely. Two script-specific marks are used; U+085D ◦ MANDAIC SMALL PUNCTUATION represents a minor break (analogous to a comma), and U+085E ◉ MANDAIC LARGE PUNCTUATION represents a major break (analogous to a full stop). In legacy fonts these are encoded on COMMA and FULL STOP respectively.

5. Collating order. The order of the letters in the code chart is their alphabetical order. The diacritical marks do not affect primary weights but are taken into consideration in tie-breaking.

6. Character names. The transliteration in the character names follows the usual UCS naming conventions, although ASZ has been preferred to what might have been written with ASS.

7. Linebreaking. Line-breaking properties for Mandaic are the same as those for Syriac. To justify text, U+085F ‒ MANDAIC KASHIDA is quite often used. In legacy fonts this is encoded on LOW LINE. The characters U+ 0640 ARABIC TATWEEL, U+180A MONGOLIAN NIRUGU, and U+ 07FA NKO LAJANYALAN suggest to us that script-specific encoding is appropriate for scripts which use extenders of this kind. Accordingly, the N'Ko, Mongolian, or Arabic characters (which have N'Ko, Mongolian, and Arabic script properties) should not be used for Mandaic (or for Manichaean, or for Psalter Pahlavi).

8. Unicode Character Properties.

```
0840;MANDAIC LETTER HALQA;Lo;0;R;;;;;N;;;;;
0841;MANDAIC LETTER AB;Lo;0;R;;;;;N;;;;;
0842;MANDAIC LETTER AG;Lo;0;R;;;;;N;;;;;
0843;MANDAIC LETTER AD;Lo;0;R;;;;;N;;;;;
0844;MANDAIC LETTER AH;Lo;0;R;;;;;N;;;;;
0845;MANDAIC LETTER USHENNA;Lo;0;R;;;;;N;;;;;
0846;MANDAIC LETTER AZ;Lo;0;R;;;;;N;;;;;
0847;MANDAIC LETTER IT;Lo;0;R;;;;;N;;;;;
0848;MANDAIC LETTER ATT;Lo;0;R;;;;;N;;;;;
0849;MANDAIC LETTER AKSA;Lo;0;R;;;;;N;;;;;
084A;MANDAIC LETTER AK;Lo;0;R;;;;;N;;;;;
084B;MANDAIC LETTER AL;Lo;0;R;;;;;N;;;;;
084C;MANDAIC LETTER AM;Lo;0;R;;;;;N;;;;;
084D;MANDAIC LETTER AN;Lo;0;R;;;;;N;;;;;
084E;MANDAIC LETTER AS;Lo;0;R;;;;;N;;;;;
084F;MANDAIC LETTER IN;Lo;0;R;;;;;N;;;;;
0850;MANDAIC LETTER AP;Lo;0;R;;;;;N;;;;;
0851;MANDAIC LETTER ASZ;Lo;0;R;;;;;N;;;;;
0852;MANDAIC LETTER AQ;Lo;0;R;;;;;N;;;;;
0853;MANDAIC LETTER AR;Lo;0;R;;;;;N;;;;;
0854;MANDAIC LETTER ASH;Lo;0;R;;;;;N;;;;;
0855;MANDAIC LETTER AT;Lo;0;R;;;;;N;;;;;
0856;MANDAIC LETTER DUSHENNA;Lo;0;R;;;;;N;;;;;
0857;MANDAIC LETTER KAD;Lo;0;R;;;;;N;;;;;
0858;MANDAIC LETTER AIN;Lo;0;R;;;;;N;;;;;
0859;MANDAIC AFFRICATION MARK;Mn;220;NSM;;;;;N;;;;;
085A;MANDAIC VOCALIZATION MARK;Mn;220;NSM;;;;;N;;;;;
085B;MANDAIC GEMINATION MARK;Mn;220;NSM;;;;;N;;;;;
085D;MANDAIC SMALL PUNCTUATION;Po;0;ON;;;;;N;;;;;
085E;MANDAIC LARGE PUNCTUATION;Po;0;ON;;;;;N;;;;;
085F;MANDAIC KASHIDA;Lm;0;R;;;;;N;;;;;
```

9. Bibliography.

Al-Mubaraki, Brayan Majid, and Majid Fandi Al-Mubaraki. 2006. *A Mandaic dictionary: Mandaic-English*. Sydney, Australia: Majid Fandi Al-Mubaraki.

Daniels, Peter T., and William Bright, eds. 1996. *The world's writing systems*. New York; Oxford: Oxford University Press. ISBN 0-19-507993-0

Faulmann, Carl. 1990 (1880). *Das Buch der Schrift*. Frankfurt am Main: Eichborn. ISBN 3-8218-1720-8

Haarmann, Harald. 1990. *Die Universalgeschichte der Schrift*. Frankfurt: Campus. ISBN 3-593-34346-0

Macuch, R, & E. S. Drower. 1963. *A Mandaic dictionary*. Oxford: Clarendon Press.

	084	085
0	o 0840	09 0850
1	0841	0851
2	0842	0852
3	0843	0853
4	0844	0854
5	0845	0855
6	| 0846	0856
7	0847	0857
8	0848	ع 0858
9	0849	0859
A	084A	085A
B	084B	085B
C	084C	
D	084D	085D
E	084E	085E
F	084F	— 085F

Letters

0840	o	MANDAIC LETTER HALQA
		a
0841		MANDAIC LETTER AB
0842		MANDAIC LETTER AG
0843		MANDAIC LETTER AD
0844		MANDAIC LETTER AH
0845		MANDAIC LETTER USHENNA
		u
0846	|	MANDAIC LETTER AZ
0847		MANDAIC LETTER IT
0848		MANDAIC LETTER ATT
0849		MANDAIC LETTER AKSA
		i
084A		MANDAIC LETTER AK
084B		MANDAIC LETTER AL
084C		MANDAIC LETTER AM
084D		MANDAIC LETTER AN
084E		MANDAIC LETTER AS
084F		MANDAIC LETTER IN
0850		MANDAIC LETTER AP
0851		MANDAIC LETTER ASZ
0852		MANDAIC LETTER AQ
0853		MANDAIC LETTER AR
0854		MANDAIC LETTER ASH
0855		MANDAIC LETTER AT
0856		MANDAIC LETTER DUSHENNA
		di
0857		MANDAIC LETTER KAD
0858	ع	MANDAIC LETTER AIN

Diacritics

0859		MANDAIC AFFRICATION MARK
085A		MANDAIC VOCALIZATION MARK
085B		MANDAIC GEMINATION MARK

Punctuation

085D		MANDAIC SMALL PUNCTUATION
085E		MANDAIC LARGE PUNCTUATION

Letter extender

085F	—	MANDAIC KASHIDA

Figures

MANDAEAN ALPHABET WITH LATIN AND
HEBREW TRANSLITERATIONS

MANDAEAN	LATIN	HEBREW
	a	א
	b	ב
	ġ	ג
	d	ד
	h	ה
	u	ו
	z	ז
	ḥ	ח
	ṭ	ט
	i	י
	k	כ
	l	ל
	m	מ
	n	נ
	s	ס
	ʿ	ע
	p	פ
	ṣ	צ
	q	ק
	r	ר
	š	ש
	t	ת
	ḏ-	ד

(as procl.,
in ᵛ as kḏ)

Figure 1. Chart with transliterations into Latin and Hebrew from Macuch 1963.
DUSHENNA and KAD are given at the end.

Figure 2. Chart showing basic syllables in Mandaic.
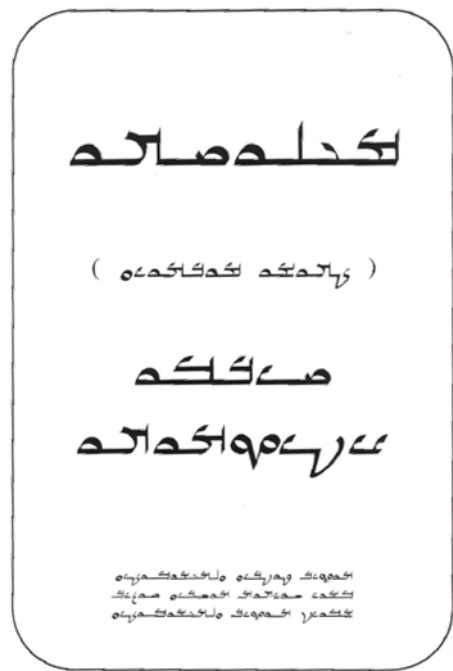
Figure 3. Sample text in Mandaic.
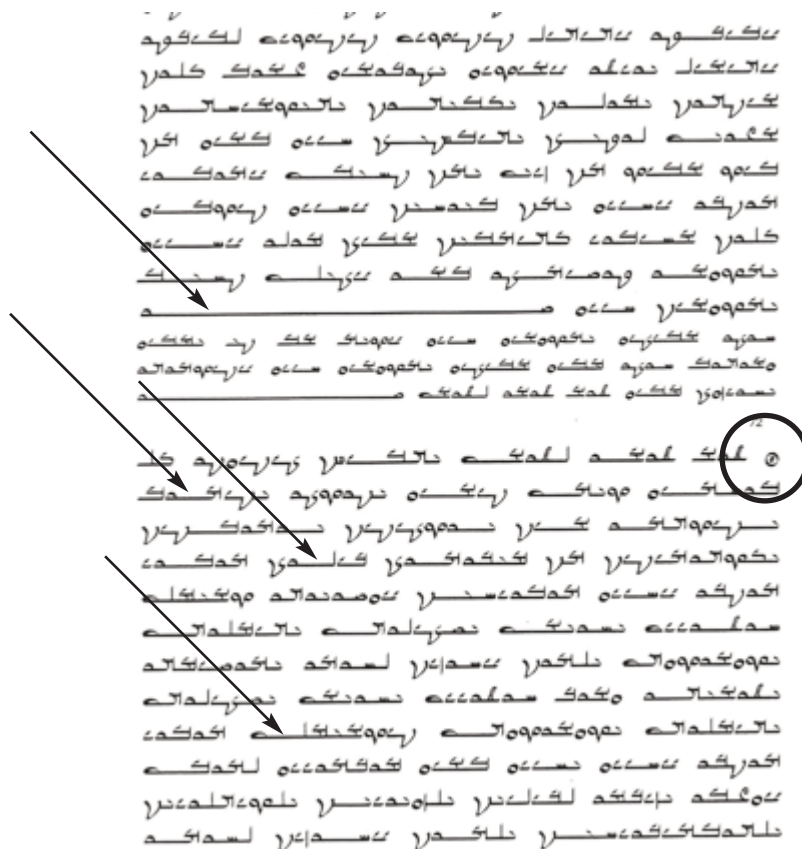
Figure 4. Sample text in Mandaic. Note the frequent use of extensions with U+085F – MANDAIC KASHIDA and the use of U+085E ⊙ MANDAIC LARGE PUNCTUATION.

Figure 5. Sample text in Mandaic. Note the use of U+085F MANDAIC KASHIDA and the frequent use U+085D ⁏ MANDAIC SMALL PUNCTUATION.

Figure 6. Sample text in Mandaic.

ḫtqapn	amla	nm	ḫnim	tqapnd	amšin	kabuṭ	kabuṭ←

aišib	rudḏ	aruadl	ḫbtiuhḏ	airas	argaplu	amuliql

aigulpu	aniq	anisḏ	akušhḏ	amlal	aiṭah	ḫlukd	artal

Figure 7. Sample of Mandaic text from Daniels 1996.

Figure 8. Sample text in Mandaic showing extended letters for foreign sounds. This is a modern letter not an old manuscript.

Figure 9. Sample text in Mandaic showing extended letters for foreign sounds. This is a modern letter not an old manuscript.

Figure 10. Sample text in Mandaic.

Figure 11. Sample text in Mandaic.

MANDÄISCH.

Isolirt	Verbunden	Wert	Isolirt	Verbunden	Wert	Isolirt	Verbunden	Wert
○	ℚ ▱	a	⅃	⅃	ṭ	⊐ ⊿	⊐ ⊿	₃
⊥	⊥ ⊿	b	∠	∠	y	∿	∿	f
⅃	⅃ ⅃	g	∿	∿	k	∿	∿	ṣ
⅃	⅃	d	⅃	⅃	l	⅃	⅃	q
⋙	⋙	h, χ	⅃	⅃	m	⊒	⊒	r
⅃	⊥	v	∿	∿	n	⅏ ⋙	⅏ ⋙	š
∫	⅃	z	⅃ ⅃	⅃ ⅃	s	⅃	⅃	t

Ligaturen.

Zeichen	Wert	Zeichen	Wert	Zeichen	Wert
⅃	di, de	⅃∿	nd	⅃∿	ṣv
⅃∿	kd	∿	nv	⅃∿	ṣl
⅃∿	kr	∿	ny	⅃∿	ṣr
∿	ky	∿	nt	⅃∿	nq
∿	kl	∿	pv	⅃	vt
∿	kv	∿	pl	⅏	š
∿	kt	⅃∿	pr		

Die Umgebungen von Wasith, Howaizah und Bassora werden noch jetzt von einem Volke bewohnt, welches sich Ṣabier, Nazaräer oder Mandäer nennt. Die Araber nennen sie auch Nabathäer, Galiläer oder St. Johannes-christen. Der Glaube dieses Volkes ist eine Mischung von christlichen Lehren mit den alten Landestraditionen. Ihre Sprache ist die aramäische und nähert sich der talmudisch-babylonischen. Die Schrift unterscheidet sich von den verwandten dadurch, dass א ו י Vokale sind, die Gutturale ח und ה, sowie א und ע nicht unterschieden werden und geschrieben, wie gesprochen wird, daher auch alle Vokale geschrieben werden; das n dient als Konsonantenverdopplung. Die Schrift zeigt alterthümliche, an das Aramäische sich lehnende Formen.

Figure 12. Chart of Mandaic letters and ligatures from Faulmann 1880. The specific glyphs of these ligatures would be font-specific; they follow the normal shaping rules as in a standard font:

⅃	dušenna	⅃∿	nd	∿	ṣu
⅃∿	kd	∿	nu	⅃∿	ṣl
⅃∿	kr	∿	ni	⅃∿	ṣr
⅃∿	ki	∿	nt	⅃∿	nq
∿	kl	∿	pu	⅃	ut
∿	ku	∿	pl	⅏	š
⅃∿	kt	⅃∿	pr		

Le Livre d'Adam

Figure 13. Sample text in Mandaic from the Imprimerie Nationale.

A. Administrative

1. Title
Proposal for encoding the Mandaic script in the BMP of the UCS
2. Requester's name
Michael Everson
3. Requester type (Member body/Liaison/Individual contribution)
Individual contribution.
4. Submission date
2008-08-04
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
Mandaic.
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
31.
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 B.1.
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?
Brian Mubaraki and Michael Everson.
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?
Brian Mubaraki, Charles Häberl, Jorunn Buckley, William Clocksin
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?

There are some 60,000-70,000 Mandaeans worldwide.

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

Traditional and liturgical use.

4b. Reference

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

Yes.

5b. If YES, where?

In Iran, Iraq, and elsewhere.

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?