Amplification and Sequencing of Full-Length HEV Genomes

Total RNA was subjected to RT-PCR for amplification of the nearly full-length 7kb sequence of the HEV genome. The RNA was reverse-transcribed with PrimeScript Reverse Transcriptase (TaKaRa Biomedicals, Shiga, Japan) and subjected to a first round of amplification using PrimeSTAR GXL DNA Polymerase (TaKaRa Biomedicals). The primers used for reverse transcription, the first round of PCR, and any second round of PCR are indicated in Fig. (**1**). Nearly full-length cDNA of $G3_{JP}$, $G3_{SP}$ and $G3_{US}$ could be generated in the first round. That of $G4_{JP}$ was generated in the second round. The 5'-end sequence was determined by an RNA ligase-mediated rapid amplification of cDNA ends (RLM-RACE) technique with the First Choice RLM-RACE kit (Ambion, Austin, TX). RNA was treated with calf intestinal alkaline phosphatase followed by tobacco acid pyrophosphatase, and then ligated to an RNA adaptor supplied in the kit. RT-PCR was performed with an OneStep RT-PCR Kit (QIAGEN, Hilden, Germany). The second round of PCR was performed using Hot Start Taq DNA Polymerase (QIAGEN). The primers for these PCRs are also shown in Fig. (**1**). The 3'-end sequence was determined with a 3'-Full RACE Core Set (TaKaRa Biomedicals) according to the directions. Primers for 3'-end amplification are shown in Fig. (**1**). The 5'-terminal and 3'-terminal PCR products were ligated with pT7Blue (Novagen, San Diego, CA) and cloned into *E. coli* DH5α (TOYOBO, Osaka, Japan).

The cDNAs of nearly full length were sequenced directly using the BigDye Terminator ver1.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) on an ABI PRISM 3100-Avant Genetic Analyzer (Applied Biosystems). The cloned 5'-end and 3'-end cDNAs were also sequenced using this kit.

Sequence Analysis of PCR Products

The amplification products were extracted from agarose gels using a QIAEX^{R}II Gel Extraction Kit (QIAGEN). Both strands of the products were sequenced using the BigDye Terminator Cycle Sequencing Ready Reaction Kit (PE Applied Biosystems). Sequence analysis was performed using Genetyx version 7.0 (Genetyx, Tokyo, Japan). Sequence alignments were generated by CLUSTAL W (version 1.8) [12]. A phylogenetic tree was constructed by the neighbor-joining method, as described previously [13], based on the entire nucleotide sequence of the HEV genome. Bootstrap values were determined on 1,000 re-samplings of the data sets [14]. The final tree was obtained using the Tree View program (version 1.6.6) [15].

Sequence Comparison

The individual sequences after alignments were compared at the nucleotide level using the "Average Difference of All pairwise comparisons" mode in the DifferencePlot (http://www.gen-info.osaka-u.ac.jp/~uhmin/study/differencePlot/index.html) with a window size of 30 and 6 slides. Similarly, amino acid sequences were compared using a window size of 1 and 1 slide. Briefly, the "Average Difference of all pairwise comparisons" calculates match rates of each window for all possible sequence pairs, and then calculates the average match rate of all the pairs for each window. Finally, a DifferencePlot displays the average match rate as two types of graph (A. Yamashita and T. Yasunaga, Personal Communication).

RESULTS

Full-Length Sequencing of Four Representative HEVs Derived from Swine Fecal Samples in Japan

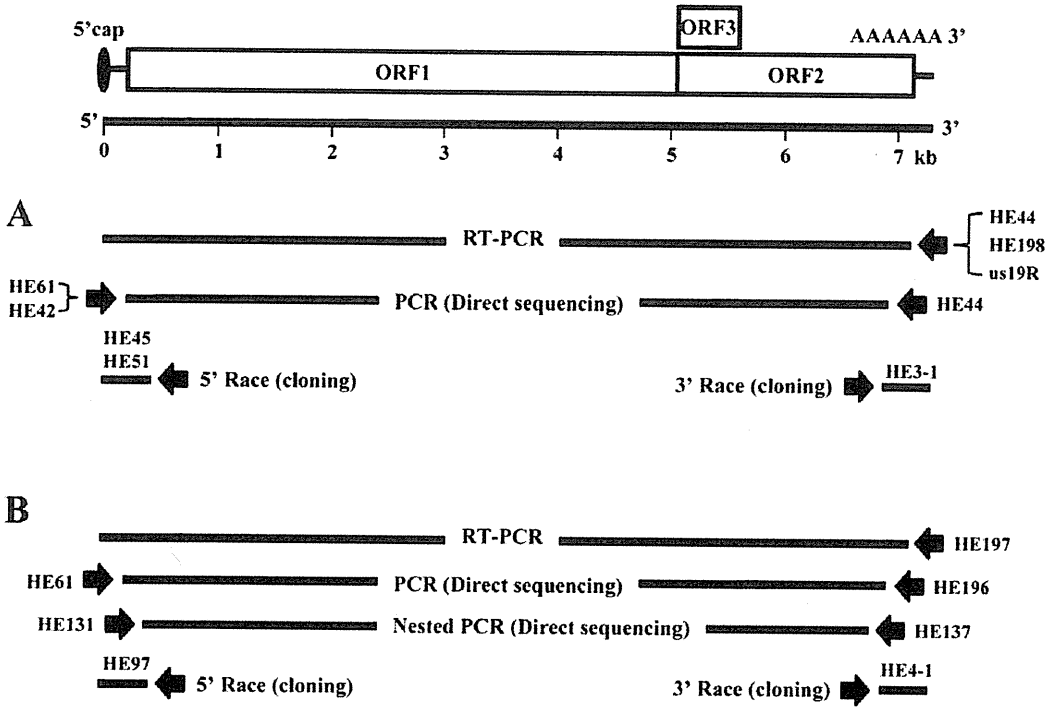
A total of 320 fecal samples were collected from 32 commercial farms in Japan. Among them, the 159 HEV positive samples were subjected to genotyping by RT-PCR with primers at ORF2 followed by sequencing [11]. All the HEV sequences, except for several unclassified types, were classified into four clusters, $G3_{JP}$, $G3_{SP}$, $G3_{US}$ and $G4_{JP}$. Therefore, we selected a representative sample from each of the clusters: swJR-P5 ($G3_{JP}$), swJB-E10 ($G3_{SP}$), swJB-M8 ($G3_{US}$) and swJB-H7 ($G4_{JP}$).

The original fecal samples used for genotyping were also used for partial purification of these four HEVs. Extracted nucleic acid samples were used to sequence the products of RT-PCR (Fig. **1**). The full-length sequences showed heterogeneous populations, i.e., swJB-E10 having nucleotide variation at 17 sites; swJR-P5 and swJB-H7 having nucleotide variation at 1 site; and swJB-M8 with no variation (data not shown). Therefore, we selected the predominant nucleotide sequences. Sequence accession numbers are shown in Fig. (**2**).

Comparison of Full-Length Sequences within Each Genotype and Cluster

Full-length genomic sequences of HEVs belonging to G1 (n=5), G2 (n=1), G3 (n=23) and G4 (n=13) that were available in the GenBank database were used for the comparison. Consistent with our previous analysis using sequences for part of ORF2 [11], a phylogenetic tree of the full-length genomic sequences showed that our HEV samples are located within the four individual clusters, $G3_{JP}$, $G3_{SP}$, $G3_{US}$ and $G4_{JP}$ (Fig. **2**).

Next, we estimated the variation at the nucleotide and amino acid levels, with all the sequences used for the phylogenetic analysis (Fig. **3A**). The variation was estimated by comparing sequences every 30 nucleotides window as well as every amino acid. The regions highly conserved (>90% identity) among G3 and G4 at the nucleotide level were located around ORF3. This result indicates that although the HEV sequences belonging to G3 formed independent sub-clusters, significant parts remained highly conserved at the nucleotide level. In contrast, the variable regions (<60% and 60-74% identity) were concentrated in the ORF1 V region in both genotypes. Although most of the regions were conserved at the amino acid level, the V region of ORF1 was highly variable in both G3 and G4. Further, the variation within clusters at the nucleotide level was examined using individual HEV isolates based on the full-length sequences used for Fig. (**2**) (Fig. **3B**). The results clearly showed for all four clusters that although the V region of ORF1 was variable, all the other regions were fairly well conserved. Comparisons among individual clusters revealed that the highly conserved regions spread over the full-length sequence were smaller in $G3_{JP}$. Thus, the individual isolates in a cluster showed some variation.



C

	Cluster	Sense primer (nucleotide position) sequence	Antisense primer (nucleotide position) sequence
RT primer	$G3_{JP}$		HE44 (7150-7127) GACTCCCGGGTTTTRCCTACCTTC
	$G3_{SP}$		HE198 (7227-7206) TTCCAGGGAGCGCGGAACGCAG
	$G3_{US}$		us19R (7208-7189) CAGAAATRAGAAATAAAGCC
	$G4_{JP}$		HE197 (7227-7206) TTCCAGGGAGCGCGAAACGCAG
First PCR	$G3_{JP}$	HE61 (7-31) CACRTATGTGGTCGAYGCCATGGAG	HE44 (7150-7127) GACTCCCGGGTTTTRCCTACCTTC
	$G3_{SP}$	HE61 (7-31) CACRTATGTGGTCGAYGCCATGGAG	HE44 (7150-7127) GACTCCCGGGTTTTRCCTACCTTC
	$G3_{US}$	HE42 (7-31) CACGTATGTGGTCGATGCCATGGAG	HE44 (7150-7127) GACTCCCGGGTTTTRCCTACCTTC
	$G4_{JP}$	HE61 (7-31) CACRTATGTGGTCGAYGCCATGGAG	HE196 (7150-7127) TACTCCCGGGTTTTACCCACCTTC
Second PCR	$G4_{JP}$	HE131 (18-47) TCGACGCCATGGAGGCCCATCAGTTCATAA	HE137 (6993-6968) GGCCAATGCAGAGTGAGGTGCAAGGACA
5'RACE	$G3_{JP}$	1st: HE45 (186-167) TGCCGGGGTTGCATCAAATT	outer primer GCTGATGGCGATGAATGAACACTG
		2nd: HE51 (131-111) GCCKRACYACCACAGCATTCG	inner primer CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG
	$G3_{SP}$	1st: HE45 (186-167) TGCCGGGGTTGCATCAAATT	outer primer GCTGATGGCGATGAATGAACACTG
		2nd: HE51 (131-111) GCCKRACYACCACAGCATTCG	inner primer CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG
	$G3_{US}$	1st: HE45 (186-167) TGCCGGGGTTGCATCAAATT	outer primer GCTGATGGCGATGAATGAACACTG
		2nd: HE51 (131-111) GCCKRACYACCACAGCATTCG	inner primer CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG
	$G4_{JP}$	1st: HE97 (130-111) CCGAACCACCACAGCATTCG	outer primer GCTGATGGCGATGAATGAACACTG
		2nd: HE97 (130-111) CCGAACCACCACAGCATTCG	inner primer CGCGGATCCGAACACTGCGTTTGCTGGCTTTGATG
3'RACE	$G3_{JP}$	HE3-1 (6681-6698) GCCACTGGTGCTCAGGCT	3-site Adaptor primer CTGATCTAGAGGTACCGGATCC
	$G3_{SP}$	HE3-1 (6681-6698) GCCACTGGTGCTCAGGCT	3-site Adaptor primer CTGATCTAGAGGTACCGGATCC
	$G3_{US}$	HE3-1 (6681-6698) GCCACTGGTGCTCAGGCT	3-site Adaptor primer CTGATCTAGAGGTACCGGATCC
	$G4_{JP}$	HE4-1 (6681-6698) GCKACTGGTGCYCAGGGG	3-site Adaptor primer CTGATCTAGAGGTACCGGATCC

Fig. (1). Location and direction of the primers used for the sequence analysis of full-length HEV genomes. A) The primers for RT-PCR amplify nearly all of $G3_{JP}$, $G3_{SP}$ and $G3_{US}$. B) $G4_{JP}$ was subjected to nested RT-PCR. For the 5'- and 3'-ends, the analysis was performed as described in Materials and Methods. C) Nucleotide sequences of primers and nucleotide positions are shown. The numbers in parentheses are the nucleotide positions of swJ570 (Accession No. 073912).

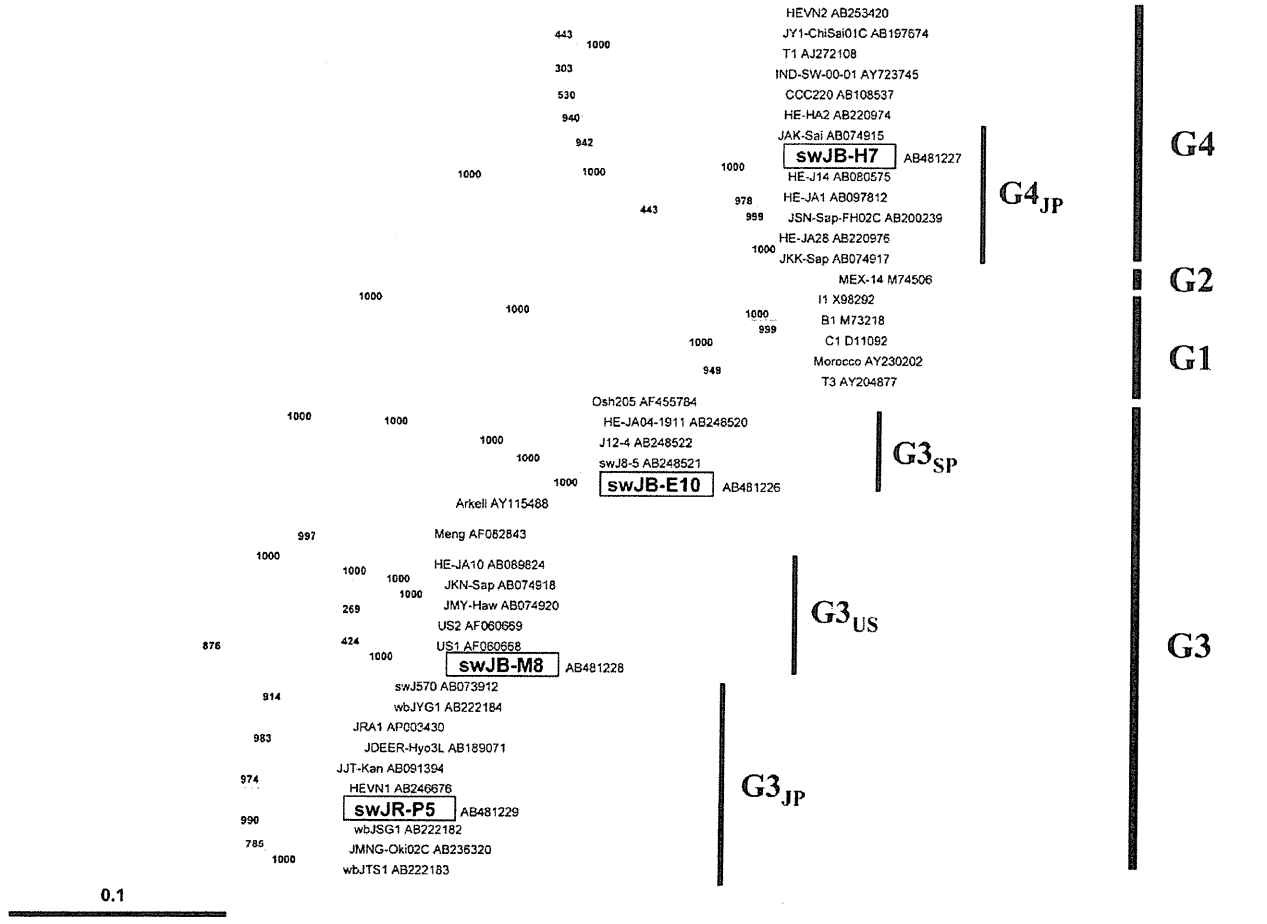


Fig. (2). Phylogenetic tree with full-length sequences of the four HEVs, together with reference HEVs. The name and accession number of individual isolates belonging to G1 to G4 are shown. The four representative HEVs in this study are boxed. The number on each branch is a bootstrap value from 1,000 re-samplings.

Comparison of Full-Length Sequences Across Clusters and Genotypes

Next, we tried a similar approach across the three clusters within G3 at the nucleotide level (Fig. **4A**). Comparisons of swJB-P5 ($G3_{JP}$) *vs* swJB-E10 ($G3_{SP}$), swJB-E10 ($G3_{SP}$) *vs* swJB-M8 ($G3_{US}$), and swJB-P5 ($G3_{JP}$) *vs* swJB-M8 ($G3_{US}$), revealed a clear distribution of variable and conserved regions over the entire sequence. Interestingly, highly conserved regions (identity in >90%) were more frequent between $G3_{JP}$ and $G3_{US}$ than the other combinations. In contrast, variable regions (<60% and 60-74% identity) were identified in the V region of ORF1, but significantly smaller between $G3_{JP}$ and $G3_{US}$. These results indicate that $G3_{JP}$ and $G3_{US}$ are more closely related than the other strains.

Further, we applied a similar approach to the cross-genotype variation between G3 and G4 (Fig. **4B**). As expected, positions of highly conserved and variable regions were consistent with those of the previous combinations shown above. However, for all three combinations [swJB-H7 ($G4_{JP}$) *vs* swJB-P5 ($G3_{JP}$), swJB-H7 ($G4_{JP}$) *vs* swJB-E10 ($G3_{SP}$), and swJB-H7 ($G4_{JP}$) *vs* sw-JB-M8 ($G3_{US}$)], the highly conserved regions (identity in >90%) were significantly narrower, whereas the variable regions were apparently wider, throughout the genomic sequence, compared with the above comparisons within genotypes (Fig. **3A**), within clusters (Fig. **3B**), and between clusters (Fig. **4A**).

As summarized in Table **1**, comparison of the full-length sequences showed similar results, i.e., $G3_{JP}$ is very similar to $G3_{US}$, somewhat different to $G3_{SP}$, and very different to $G4_{JP}$ at the nucleotide level. This tendency was also observed at the amino acid level in ORF1 and ORF3. However, for ORF2, $G3_{JP}$ was very similar to $G4_{JP}$ as well as $G3_{US}$ and $G3_{SP}$.

Variation in the Sequences for the Primers and Probes Used for Detection of the HEV Genome by RT-PCR

The comparison among the four clusters in G3 and G4 showed that the overlapping region of ORF2 and ORF3 was

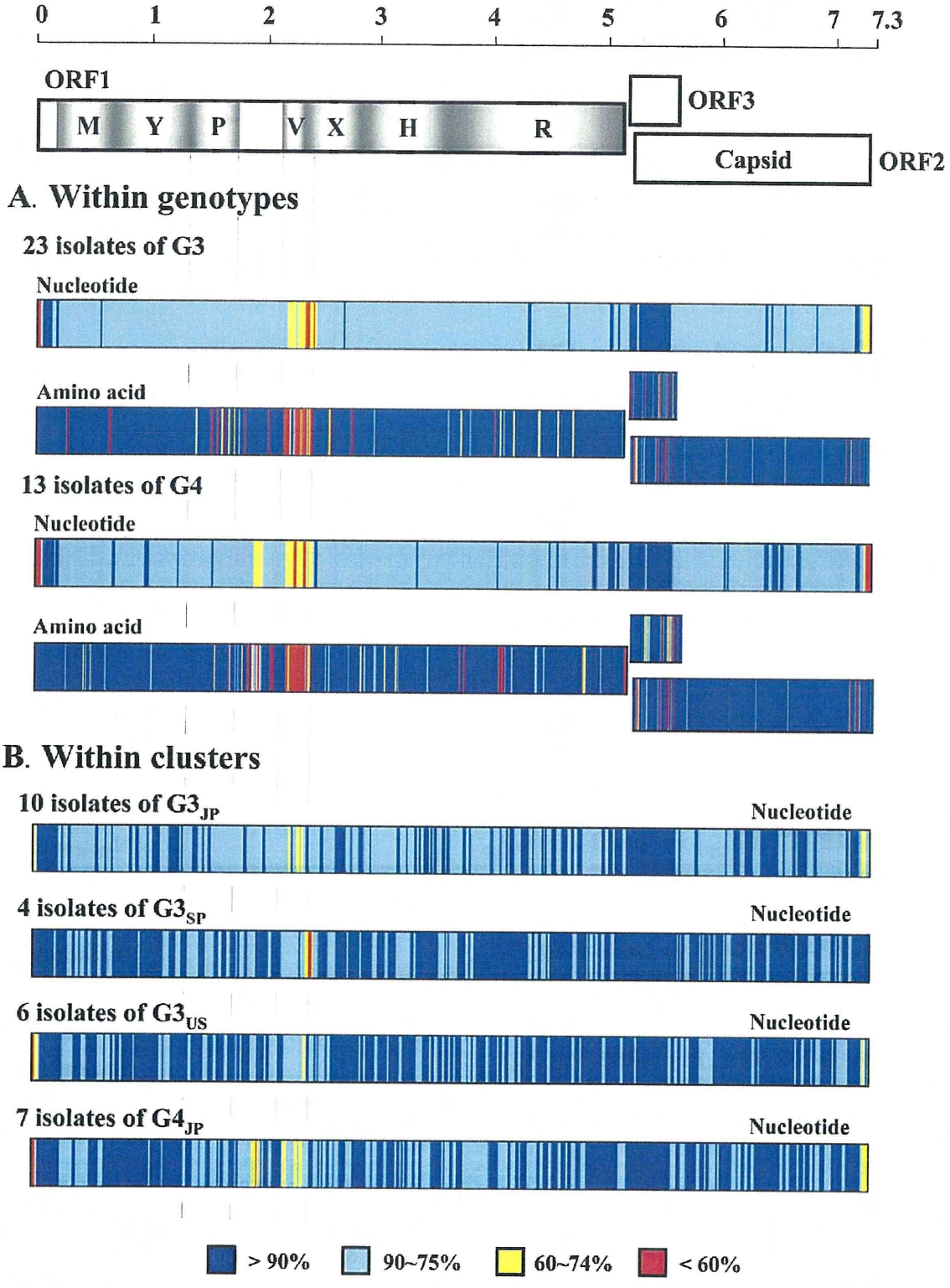
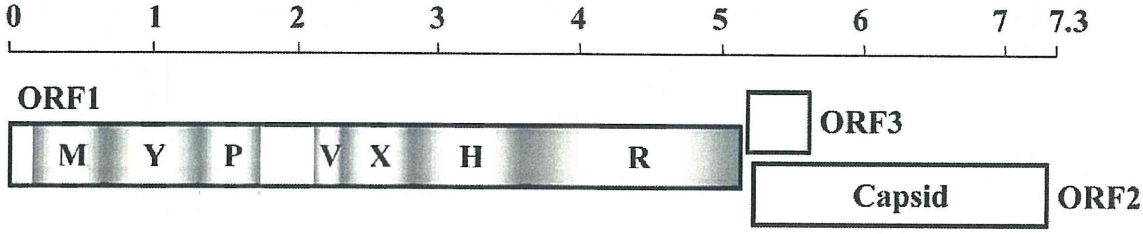


Fig. (3). Conservation within genotypes and clusters. A) A total of 23 isolates belonging to G3, including three HEVs from this study, and a total of 13 isolates belonging to G4 including one HEV from this study, were used. For the sequence comparison, the average rate of conservation (%) every 30 nucleotides or every amino acid was calculated and values are shown with different colors. B) Ten isolates including swJR-P5 in the $G3_{JP}$ cluster, 4 isolates including swJB-E10 in the $G3_{SP}$ cluster, 6 isolates including swJB-M8 in the $G3_{US}$ cluster and 7 isolates including swJB-H7 in the $G4_{JP}$ cluster were used.



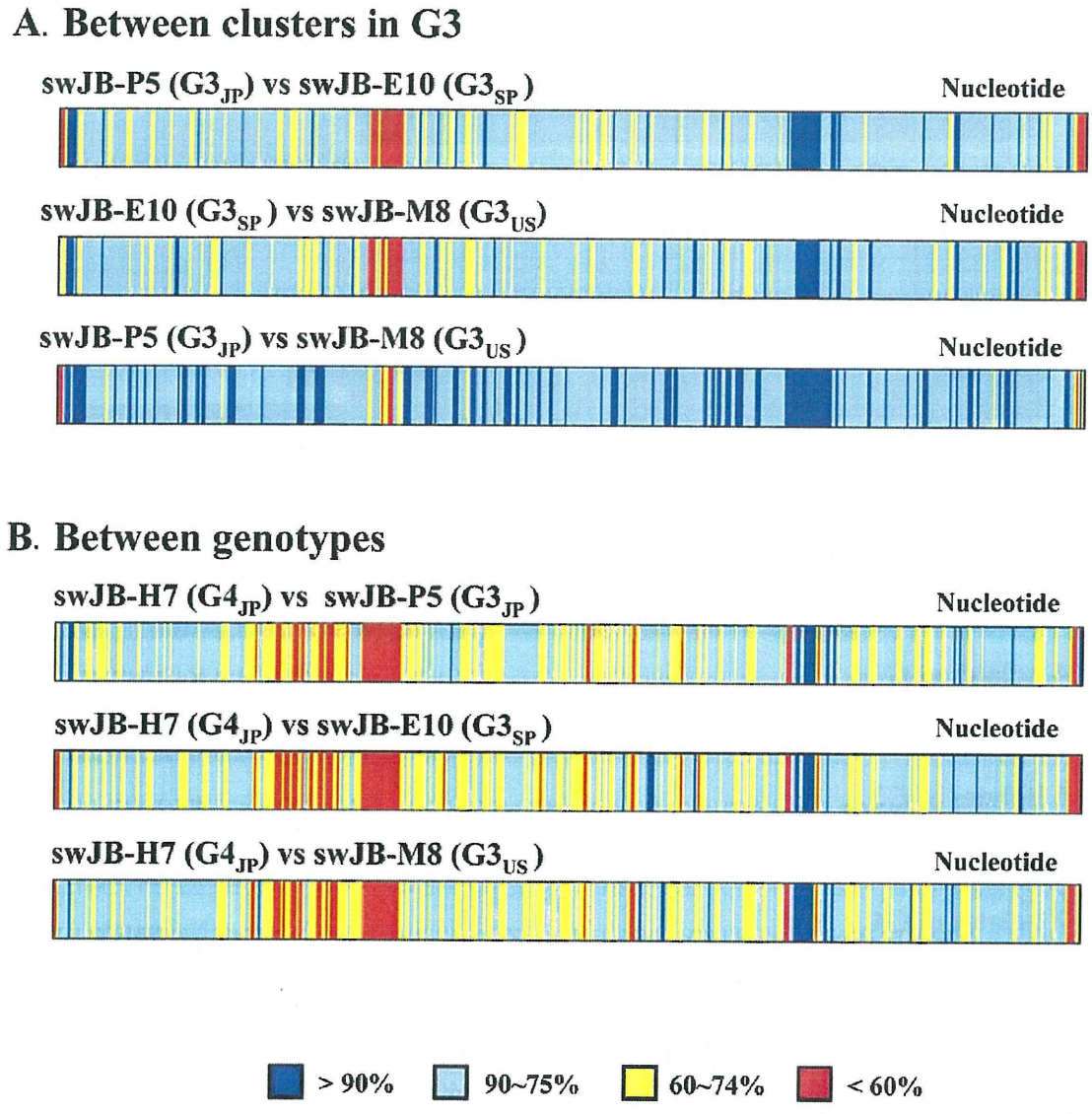


Fig. (4). Conservation across clusters in G3 and between G3 and G4 at the nucleic acid level. Two of the four isolates in this study, $G3_{JP}$, $G3_{SP}$, $G3_{US}$ and $G4_{JP}$, were compared. Sequences with different conservation rates are shown in different colors, as in Fig. (**3**).

Table 1. Homology of Individual ORF Sequences of Four HEV Isolates

% Identity in Nucleotides; Full-Length

	swJR-P5 ($G3_{JP}$)	swJB-M8 ($G3_{US}$)	swJB-E10 ($G3_{SP}$)
swJB-M8 ($G3_{US}$)	87.4		
swJB-E10 ($G3_{SP}$)	81.7	81.3	
swJB-H7 ($G4_{JP}$)	75.6	77.4	75.5

% Identity in Amino Acids; ORF1

	swJR-P5 ($G3_{JP}$)	swJB-M8 ($G3_{US}$)	swJB-E10 ($G3_{SP}$)
swJB-M8 ($G3_{US}$)	96.8		
swJB-E10 ($G3_{SP}$)	93.9	93.6	
swJB-H7 ($G4_{JP}$)	85.7	85.3	85.7

% Identity in Amino Acids; ORF2

	swJR-P5 ($G3_{JP}$)	swJB-M8 ($G3_{US}$)	swJB-E10 ($G3_{SP}$)
swJB-M8 ($G3_{US}$)	98.5		
swJB-E10 ($G3_{SP}$)	97.4	97.3	
swJB-H7 ($G4_{JP}$)	91.8	91.5	92.1

% Identity in Amino Acids; ORF3

	swJR-P5 ($G3_{JP}$)	swJB-M8 ($G3_{US}$)	swJB-E10 ($G3_{SP}$)
swJB-M8 ($G3_{US}$)	96.5		
swJB-E10 ($G3_{SP}$)	93.8	92.0	
swJB-H7 ($G4_{JP}$)	82.5	83.3	82.5

highly conserved. This makes the region a potential target for detecting HEV RNA by RT-PCR. Therefore, we next focused on the appropriateness for this region as a target for detecting HEV RNA. Several groups have performed amplification in the overlapping region of ORF2 and ORF3 [16-19]. We quantified the amount of HEV RNA according to the method of Jothikumar *et al.* [17]. Sequences from the GenBank database were used for the examination: 72 sequences from animals including pigs and 107 sequences from humans at HE86 and HE87; and 72 sequences from animals including pigs and the same 107 sequences from humans at FHE88, as shown in Table **2**. HE86, FHE88 and HE87 correspond to JVHEVF, JVHEVP and JVHEVR, respectively, in Jothikumar *et al.* [17]. Some 96.3%, 94.4% and 96.3% of the human-derived HEV sequences were conserved in HE86, FHE88 and HE87, respectively. Slightly less conservation, except for the region at HE87, was observed in the sequences from animals: 91.7%, 88.9% and 97.2% in HE86, FHE88 and HE87, respectively.

DISCUSSION

The transmission of HEV to humans in Japan is mostly food-borne [1-5]. Full-length sequences of four HEVs in fecal samples from pig farms in Japan were analyzed in this study. The four representatives were located in independent clusters in a phylogenetic tree constructed with these sequences together with full-length reference sequences from the GenBank database: $G3_{JP}$, $G3_{SP}$, $G3_{US}$ and $G4_{JP}$ covered most of the HEVs distributed among swine as well as humans in Japan. A comparison of these four sequences with the reference sequences revealed highly conserved regions that overlap between ORF2 and ORF3, as well as variable regions in the ORF1 V region. Positions of the variable as well as conserved regions were virtually conserved across clusters and genotypes, as well as within clusters and genotypes.

Comparisons between clusters revealed higher genetic similarity between $G3_{JP}$ and $G3_{US}$, which were slightly more distant from $G3_{SP}$. We also found that the regions highly conserved among G3 and G4 occur within smaller areas in the overlapping region of ORF2 and ORF3. In addition, not only V but also P regions of ORF1 were highly variable. Interestingly, the ORF1 P region is unlikely to code for a papain-like protease [20], because only a part of the HEV sequence can encode the protease (see the sequences listed by their accession numbers in Fig. **2**).

Finally, we examined how efficiently the primer and probe sets could amplify HEV RNA. The $G3_{SP}$ and $G3_{US}$ HEVs arrived in Japan through the importation of pigs from the United Kingdom [21] and United States [8], respectively. $G3_{SP}$ and $G3_{US}$, in addition to $G3_{JP}$ and $G4_{JP}$, are currently distributed in pigs and humans in Japan. These HEV clusters were highly conserved at the amino acid level, but highly variable at the nucleotide level. Therefore, a sensitive method of detecting the genome is urgently needed. Several groups have used primer and probe sets at similar sites [17-19]. The results suggest that most of the sequences from GenBank (>88%) that were derived from humans and animals would be amplified with the primers HE86 (JVHEVF)/HE87 (JVHEVP) and probe FHE88 (JVHEVR). $G3_{US}$ differs by one nucleotide from the other three isolates in the region of the probe. The sequence of probe FHE100 for $G3_{US}$ was found in only 0.9% and 2.8% of the sequences from humans and animals, respectively. Based on our experience with the detection of $G3_{US}$, FHE100 was necessary for the detection of $G3_{US}$, because FHE88 was more than 10-times less sensitive than FHE100. Thus, most of the heterogeneous genomic sequences among HEVs in humans in Japan would be covered by these primer and probe sets, although this needs to be confirmed with experimental analyses. These four HEVs could be useful as a standard for heterogeneous HEV genomic sequences.

ACKNOWLEDGEMENTS

We thank Dr. Hiroshi Yasue, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan for valuable discussions. This work was conducted based on collaborative research projects between Osaka University and Benesis Corporation and between Rakuno Gakuen University and Benesis Corporation, and supported in part by a grant from the Japan Human Sciences Foundation (Grant number KHB1011).

Table 2. Nucleotide Variation in Regions Used for Primer and Probe Sets to Detect the HEV Genome with High Sensitivity

Sense Primer

Virus Origin	Nucleotide Position 5286 -5303 of swJ570 (AB073912)	Sequence Number	Appearance Ratio (%)	Primer Name
Swine, Wild boar, Deer and others, including of unknown origin	GGT GGT TTC TGG GGT GAC	66	91.7 %	HE86 (JVHEVF)
	A..	6	8.3 %	
Human	GGT GGT TTC TGG GGT GAC	103	96.3 %	HE86 (JVHEVF)
	A..	3	2.8 %	
	... ... A.. ...	1	0.9 %	

Probe

Virus Origin	Nucleotide Position 5309 –5326 of swJ570 (AB073912)	Sequence Number	Appearance Ratio (%)	Probe Name
Swine, Wild boar, Deer and others, including of unknown origin	TGA TTC TCA GCC CTT CGC	64	88.9 %	FHE88 (JVHEVP)
	... ... C..	2	2.8 %	FHE100
	... C..	2	2.8 %	
	... ... T..	1	1.4 %	
	... ... A..	1	1.4 %	
	... .C	1	1.4 %	
	C.. C..	1	1.4 %	
Human	TGA TTC TCA GCC CTT CGC	101	94.4 %	FHE88 (JVHEVP)
	... ... C..	1	0.9 %	FHE100
	... ... A..	1	0.9 %	
	... ... .T.	1	0.9 %	
	... ... .C. ...	1	0.9 %	
	... .Y. ... ... Y = T /C	1	0.9 %	
	... ...T	1	0.9 %	

Anti-Sense Primer

Virus Origin	Nucleotide Position 5338 –5355 of swJ570 (AB073912)	Sequence Number	Appearance Ratio (%)	Primer Name
Swine, Wild boar, Deer and others, including of unknown origin	TTC ATC CAA CCA ACC CCT	70	97.2 %	HE87 (JVHEVR)
	... ...T ...	1	1.4 %	
	.C.	1	1.4 %	
Human	TTC ATC CAA CCA ACC CCT	103	96.3 %	HE87(JVHEVR)
	... ...T ...	3	2.8 %	
	... ...T	1	0.9 %	

Probe FHE100, differing by one nucleotide from FHE88, was designed to detect swJB-M8 ($G3_{US}$).

REFERENCES

[1] Li TC, Chijiwa K, Sera N, Ishibashi T, *et al.* Hepatitis E virus transmission from wild boar meat. Emerg Infect Dis 2005; 11: 1958-60.

[2] Matsuda H, Okada K, Takahashi K, *et al.* Severe hepatitis E virus infection after ingestion of uncooked liver from a wild boar. J Infect Dis 2003; 188: 944.

[3] Tamada Y, Yano K, Yatsuhashi H, *et al.* Consumption of wild boar linked to cases of hepatitis E. J Hepatol 2004; 40: 869-70.

[4] Tei S, Kitajima N, Takahashi K, *et al.* Zoonotic transmission of hepatitis E virus from deer to human beings. Lancet 2003; 362: 371-3.

[5] Yazaki Y, Mizuo H, Takahashi M, *et al.* Sporadic acute or fulminant hepatitis E in Hokkaido, Japan, may be food-borne, as suggested by the presence of hepatitis E virus in pig liver as food. J Gen Virol 2003; 84: 2351-7.

[6] Emerson SU, Anderson D, Arankalle A, *et al.* In: Fauquet CM, Ed. Hepevirus. In: Virus taxonomy: eighth report of the international committee on taxonomy of viruses. Elsevier/Academic Press, London, 2004; 851-855.

[7] Schlauder GG, Mushahwar IK. Genetic heterogeneity of hepatitis E virus. J Med Virol 2001; 65: 282-92.

[8] Takahashi M, Nishizawa T, Miyajima H, *et al.* Swine hepatitis E virus strains in Japan form four phylogenetic clusters comparable with those of Japanese isolates of human hepatitis E virus. J Gen Virol 2003; 84: 851-62.

[9] Takahashi M, Nishizawa T, Tanaka T, *et al.* Correlation between positivity for immunoglobulin A antibodies and viraemia of swine hepatitis E virus observed among farm pigs in Japan. J Gen Virol 2005; 86: 1807-13.

[10] Lu L, Li C, Hagedorn CH. Phylogenetic analysis of global hepatitis E virus sequences: genetic diversity, subtypes and zoonosis. Rev Med Virol 2006; 16: 5-36.

[11] Sapsutthipas S, Urayama T, Yamate M, *et al.* Sequence variation in hepatitis E virus genotypes 3 and 4 from swine fecal samples in Japan. Open Vet Sci J 2009; 3: 68-75.

[12] Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994; 22: 4673-80.

[13] Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 1987; 4: 406-25.

[14] Felsenstein J. Confidence limits on phylogenies: An approach using the bootstrap. Evolution 1985; 39: 783-91.

[15] Page RD. TreeView: an application to display phylogenetic trees on personal computers. Comput Appl Biosci 1996; 12: 357-8.

[16] Enouf V, Dos Reis G, Guthmann JP, *et al.* Validation of single real-time TaqMan PCR assay for the detection and quantitation of four major genotypes of hepatitis E virus in clinical specimens. J Med Virol 2006; 78: 1076-82.

[17] Jothikumar N, Cromeans TL, Robertson BH, *et al.* A broadly reactive one-step real-time RT-PCR assay for rapid and sensitive detection of hepatitis E virus. J Virol Methods 2006; 131: 65-71.

[18] Orrù G, Masia G, Orrù G, *et al.* Detection and quantitation of hepatitis E virus in human faeces by real-time quantitative PCR. J Virol Methods 2004; 118: 77-82.

[19] Zhao C, Li Z, Yan B, *et al.* Comparison of real-time fluorescent RT-PCR and conventional RT-PCR for the detection of hepatitis E virus genotypes prevalent in China. J Med Virol 2007; 79: 1966-73.

[20] Koonin EV, Gorbalenya AE, Purdy MA, *et al.* Computer-assisted assignment of functional domains in the nonstructural polyprotein of hepatitis E virus: delineation of an additional group of positive-strand RNA plant and animal viruses. Proc Natl Acad Sci USA 1992; 89: 8259-63.

[21] Inoue J, Takahashi M, Ito K, *et al.* Analysis of human and swine hepatitis E virus (HEV) isolates of genotype 3 in Japan that are only 81-83 % similar to reported HEV isolates of the same genotype over the entire genome. J Gen Virol 2006; 87: 2363-9.

Received: November 24, 2009 Revised: February 25, 2010 Accepted: April 28, 2010

ARTICLES

【新連載】 バイオ医薬品の品質・安全性評価シリーズ（第１回）

バイオ医薬品の物理的化学的性質解析の現状

Current state in analysis of physicochemical properties of biopharmaceuticals

国立医薬品食品衛生研究所 生物薬品部
橋井則貴，原園 景，川崎ナナ
NORITAKA HASHII, AKIRA HARAZONO, NANA KAWASAKI
Division of Biological Chemistry and Biologicals, National Institute of Health Sciences

はじめに

バイオ医薬品は，糖尿病，肝炎，およびある種の血液関連疾患やがん治療などにおける標準的治療薬となっており，医療上の重要性は増す一方である。また，標準的治療法のない疾患への新たな治療法の提供，患者のQOLの向上，および医薬品産業の活性化を目標として，さらなる新規なバイオ医薬品の開発が進んでいる。新規なバイオ医薬品の早期実用化には，品質・安全性確保のための評価科学の進展が不可欠である。本シリーズでは，バイオ医薬品の品質・安全性評価に関する現状を，物理的化学的性質解析，生物学的性質解析・免疫化学的性質解析，不純物解析，外来性感染性物質安全性評価，製造工程評価，規格及び試験方法，免疫原性評価，および安定性評価等に分けて紹介する。

第１回は，バイオ医薬品の物理的化学的性質解析のための最新技術を，分析例を紹介しながら概説する。

1．物理的化学的性質解析

バイオ医薬品の有効性・安全性を確保するために，品質，すなわち，特性，製造方法，規格及び試験方法，安定性の評価は不可欠である。中でも特性は，製造工程開発／管理手法の開発や規格及び試験方法の設定の根拠となるものであり，開発段階では可能な限り広範囲かつ詳細に解析しなければならない。バイオ医薬品開発において明らかにすることが求められる特性は，構造・組成，物理的化学的性質，生物学的性質，免疫化学的性質，および不純物などである。構造・組成と物理的化学的性質を明確に区別することは難しいが，「生物薬品（バイオテクノロジー応用医薬品／生物起源由来医薬品）の規格及び試験方法の設定について（平成13年５月１日，医薬審発第571号厚生労働省医薬局審査管理課長通知）」では，解析すべき構造・組成として一次構造や糖鎖などの構造が，物理的化学的性質については，分子量・分子サイズ，

表１　物理的化学的性質解析の項目および主な分析法

特性解析	項目		主な分析法	
物理的化学的性質	分子量・分子サイズ		MS，HPLC，超遠心分析法，SDS-PAGE	
	不均一性	目的物質	グリコフォーム	IEF，CE，HPLC，MS
			意図的修飾	HPLC，CE，LC/MS，LC/MS/MS
		分子変化体	ジスルフィド結合ミスマッチ体	LC/MS，LC/MS/MS，HPLC，SDS-PAGEペプチドマッピング
			酸化体	LC/MS，LC/MS/MS
			脱アミド体およびアスパラギン酸の異性体	HPLC，IEF，LC/MS，LC/MS/MS
	高次構造解析	分光学的性質	NMR，X線結晶構造解析法，CD，FTIR	
		構造安定性	DSC	

MS，質量分析法；HPLC，高速液体クロマトグラフィー法；SDS-PAGE，ドデシル硫酸ナトリウム－ポリアクリルアミドゲル電気泳動法；IEF，等電点電気泳動法；CE，キャピラリー電気泳動法；LC/MS，液体クロマトグラフィー／質量分析法；LC/MS/MS，液体クロマトグラフィー／タンデム質量分析法；NMR，核磁気共鳴法；CD，円偏光二色性測定法；FTIR，フーリエ変換赤外吸収スペクトル測定法；DSC，示差走査カロリメトリー法

不均一性を示す性質として高速液体クロマトグラフィー（HPLC）パターン，アイソフォームパターンおよび電気泳動パターン，ならびに高次構造等を示す性質として分光学的性質等が記されている（**表1**）。一次構造や糖鎖などの解析手法はほぼ確立されているといえるかもしれないが，物理的化学的性質を解析するための新たな手法の開発は課題である。

2．分子量・分子サイズ

分子量，分子サイズを解析する方法として，超遠心分析法，サイズ排除クロマトグラフィー法，ドデシル硫酸ナトリウム－ポリアクリルアミドゲル電気泳動（SDS-PAGE）法などが用いられてきた。現在では，質量分析（MS）法により，質量を測定することが多い。MSによりペプチドおよびタンパク質医薬品の質量を測定するときに留意すべき点は，第十六改正日本薬局方参考情報[1]にまとめられている。

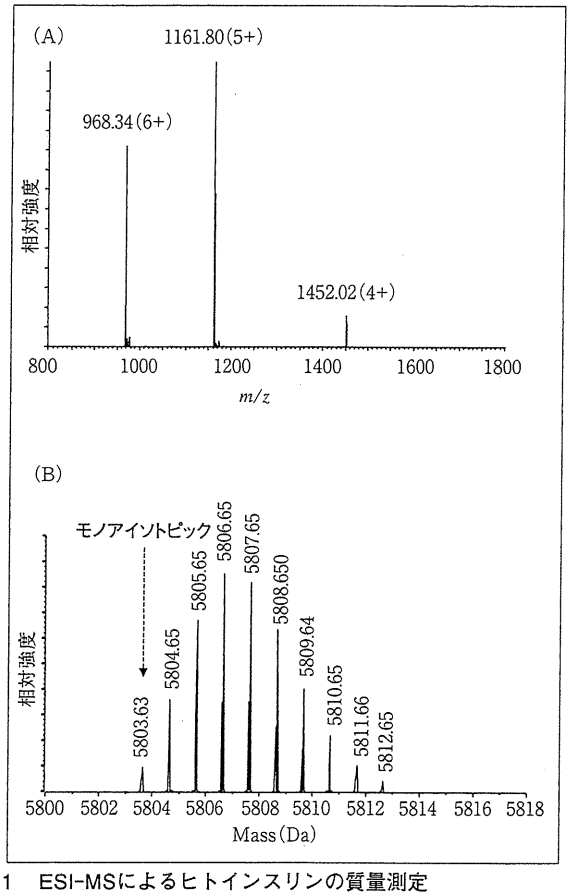


図1　ESI-MSによるヒトインスリンの質量測定
（A）マススペクトル，（B）デコンボリューションマススペクトル
MS装置，Qstar elite（Applied Biosystems），*m/z*範囲，800～2,000

図1（A）は，エレクトロスプレーイオン化（ESI）-MSにより得られたヒトインスリンのマススペクトルである。ヒトインスリンは4～6価のイオンとして検出されている。図1（B）は，（A）のデコンボリューションマススペクトルであり，モノアイソトピックピークから質量（5803.6Da）を求めることができる。このようにMSは，従来の方法よりも，分子量に近い値を求めることができる。

3．不均一性

バイオ医薬品は，細胞等の生合成過程を生産に利用しているため，分子構造上，不均一な分子種の集合体として産生される。これらの中で，DNA塩基配列から予想されるアミノ酸配列を有し，かつ適切な翻訳後修飾を持つ分子種は「目的物質」とよばれる。目的物質であっても，不均一性をもつことがある。その代表例に，糖鎖の構造や結合位置の違いにより生じるグリコフォームがある。特性解析においては，糖鎖構造と活性等の関係を考慮しながら，適切な手法を用いてグリコフォームパターンを確認する必要がある。例えば，シアル酸結合医薬品の特性解析では，シアル酸結合数がタンパク質の血中半減期に影響することがあるので，シアル酸結合数に由来する不均一性プロファイルを測定し，薬理プロファイルとの相関性を調べることが望ましい。不均一性プロファイルを得る方法として，MS，等電点電気泳動法，キャピラリー電気泳動法，および陰イオン交換クロマトグラフィー法などがある。

図2に，ESI-MSによって，チャイニーズハムスター卵巣（CHO）細胞産生ヒトエリスロポエチン（EPO）のグリコフォームを解析した例を示す。EPOは165個のアミノ酸からなる糖タンパク質であり，3カ所のN結合型糖鎖結合部位（Asn24，38および83）と1カ所のO結合型糖鎖結合部位（Ser126）を有する。結合糖鎖の大部分はシアル酸付加糖鎖であり，そのシアル酸付加糖鎖の分布が活性に影響することが知られている[2]。図2（A）はマススペクトルで，EPOは10～16価のイオンとして検出されている。図2（B）はデコンボリューションマススペクトルで，グリコフォームプロファイルと見なすことができるだろう。MSを用いたEPOのグリコフォーム解析より，①シアル酸付加数は11～14個であり，特に12および13個結合した分子種が多いこと，②アセチル化されたシアル

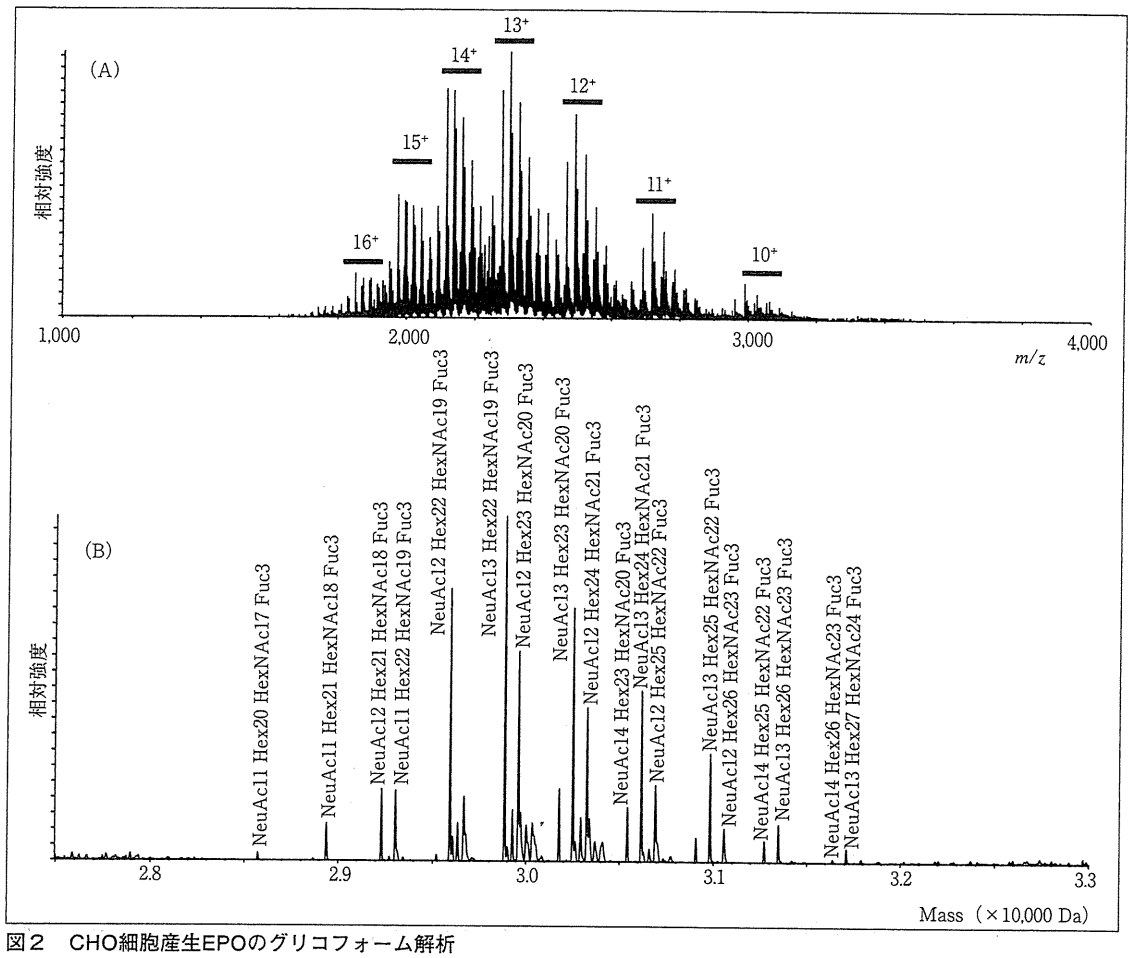


図2　CHO細胞産生EPOのグリコフォーム解析
(A)マススペクトル，(B)デコンボリューションマススペクトル
NeuAc，N-アセチルノイラミン酸；Hex，ヘキソース；HexNAc，N-アセチルヘキソサミン；Fuc，フコース。
MS装置，Qstar elite(Applied Biosystems)；*m/z*範囲，1,000～4,000。

酸がみられないこと，などがわかる。これまで，シアル酸結合タンパク質の特性解析では，負電荷分子を分離する分析手法（等電点電気泳動，キャピラリー電気泳動など）を用いてシアル酸結合数の異なるグリコフォームのパターンを解析することが多かったが，MSでは，シアル酸付加程度に加えて，糖組成，アセチル化の程度，および非ヒト糖鎖抗原であるN-グリコリルノイラミン酸の付加程度なども確認することが可能であり，今後，利用されるケースが増えるものと予想される。

目的物質が不均一性を有する糖鎖以外の例として，意図的修飾を施した医薬品があげられる。PEG化タンパク質や薬物結合抗体医薬品では，修飾分子の結合位置や結合数の違いによる不均一性が生じる。このような意図的修飾によって生じた不均一性は，キャピラリー電気泳動法，HPLC，MSなどにより解析されることが多いが，修飾分子の特徴に合った解析手法を選択することが重要である[3,4]。

バイオ医薬品の製造工程では，目的物質以外に，脱アミド体，異性体，ジスルフィド結合ミスマッチ体，酸化体，あるいは加水分解酵素などで断片化された切断体など，さまざまな「分子変化体」が生じる可能性がある。分子変化体のうち，生物活性，有効性および安全性の点で目的物質のそれに匹敵する性質を持つものは「目的物質関連物質」，また，生物活性，有効性および安全性の観点から目的物質に匹敵する特性を持たないものは「目的物質由来不純物」とよばれ区別される。分子不均一性解析では，目的物質，目的物質関連物質および目的物質由来不純物を区別すること，それらのプロファイルを求めておくことが重要である。抗体やリソソーム酵素製品のように末端のアミノ酸残基が欠失している分子種が混在する場合には，末端アミノ酸配列解析を行うとともに，イオン交換クロマトグラフィーやMSなどにより不均一

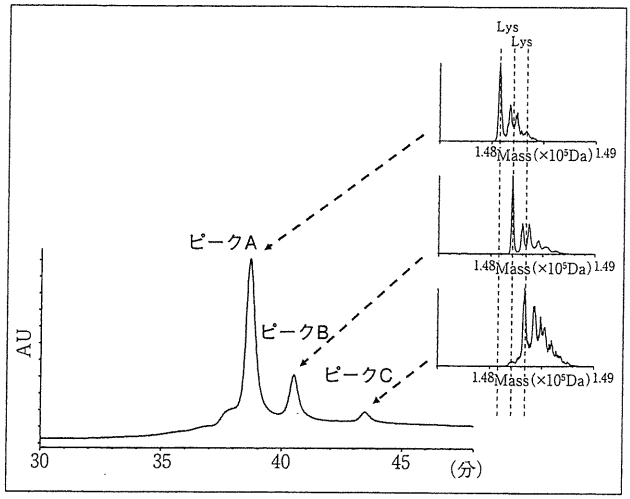


図3　CHO細胞産生ヒト抗体の不均一性解析
カラム，Agilent Bio MAb（0.1×50mm，1.7mm）；流速，0.8mL/分；UV，280nm；A溶媒，10mMリン酸ナトリウム水溶液（pH6.6）；B溶媒，500mM塩化ナトリウムを含む10mMリン酸ナトリウム水溶液（pH6.6）；グラジエント条件，5分間3%Bで流した後，50分間かけて20%Bへ上昇させた。

性プロファイルを測定することが望ましい（目的物質由来不純物の解析手法については，第3回で紹介する）。

分子変化体の解析例として，弱酸性陽イオン交換クロマトグラフィーにより，CHO細胞産生ヒト抗体のH鎖のC末端リシン（Lys）残基の欠失に由来する不均一性を解析した例を示す（**図3**）。ここでは検出されたピークA～Cを分取した後，MSにより質量を確認することで，C末端Lys残基の個数を確認している（ピークA，Lys残基をもたない分子；ピークB，Lys残基を1分子もつ分子；ピークC，Lys残基を2分子もつ分子）。また，ピーク面積比から，C末端Lys残基結合数の異なる分子の分布を確認している（ピークA，72%；ピークB，20%；ピークC，8%）。

4．分光学的性質

分光学的性質は，目的物質の高次構造情報を得るために解析される。主な分析法として，核磁気共鳴法（NMR），X線結晶構造解析法，円偏光二色性測定法（CD）およびフーリエ変換赤外吸収スペクトル測定法（FTIR）などが用いられている。

NMRの測定法のうち，バイオ医薬品の高次構造解析に用いられる測定法は，核オーバーハウザー効果（NOE）測定法であり，水素結合の位置や原子間の空間的距離を明らかにすることができる[5]。NMRは，生体内の環境に近い状態で試料を測定できる点で優れているが，現在のNMR装置で測定可能なタンパク質は，分子量3万程度とされている。一方，X線結晶構造解析法は，タンパク質の分子量に関係なく測定できることから高次構造解析に汎用されているが，糖タンパク質の結晶化が難しいなどの技術的な課題や，必ずしも生体内における高次構造を反映していない場合があることに留意する必要がある。タンパク質中の二次構造（αヘリックスやβシート）の有無についてはCDやFTIRにより解析されることが多い。

5．構造安定性

タンパク質の構造安定性を解析する手法として，示差走査カロリメトリー法（DSC）が知られている。DSCによりタンパク質の熱変性に伴う熱変化を測定することで，タンパク質の構造安定性を評価することができる。タンパク質の構造安定性は，溶媒，pH，イオン強度および添加剤等により変化するので，バイオ医薬品の開発において，DSCは製剤処方の最適化等に用いられることが多い。

バイオ医薬品の中には，サブユニットが会合し，多量体を形成している品目も多いが，多量体安定性を解析するための技術開発は発展途上である。そのため，目的の多量体を形成しているか否か，その生物活性から推測されているのが現状である。多量体安定性が有効性・安全性に直接的に関係するバイオ医薬品の1つにインスリン製剤がある。遺伝子組換えヒトインスリン製剤は，製剤中では主に六量体を形成しており，投与後に皮下組織で拡散・希釈されることで，六量体から二量体を経て単量体に解離し，吸収された後に作用を発現する。多量体の安定性はインスリン単量体が血中に移行するまでの時間に影響を与えるものと考えられており，アミノ酸残基を置換して多量体安定性を変化させて作用発現時間や持続時間を調節した遺伝子組換えインスリンアナログ製剤が次々と開発されている。インスリンスペースリスプロ（リスプロ）は超速効型インスリンアナログであり，ヒトインスリンB鎖の28番目（B28）のProとB29のLys残基が置換されている。同じ超速効型に分類されるインスリングルリジン（グルリジン）は，B3のAsnおよびB29のLys残基がそれぞれLysおよびGlu残基に置換されている。これらの置換により多量体安定性が低下し単量体に解離

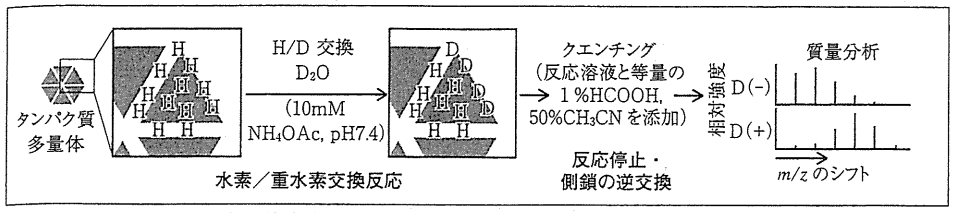


図4　HDX/MSによる多量体安定性解析の概要
重水中でタンパク質中の主鎖アミド水素原子(H)は，徐々に重水素(D)に置換される(HDX反応)。このとき，溶媒との接触の多い水素原子ほど速やかに置換される。交換された水素の数は質量の変化として質量分析計で検出される。交換反応性が特に高い官能基(-OH，-NH_2，-COOH，-SH，-$CONH_2$など)に取り込まれたDは，反応停止過程で加えられたH^+により元に戻るため，最終的に検出されるのは主に主鎖のアミド基に取り込まれたDである。

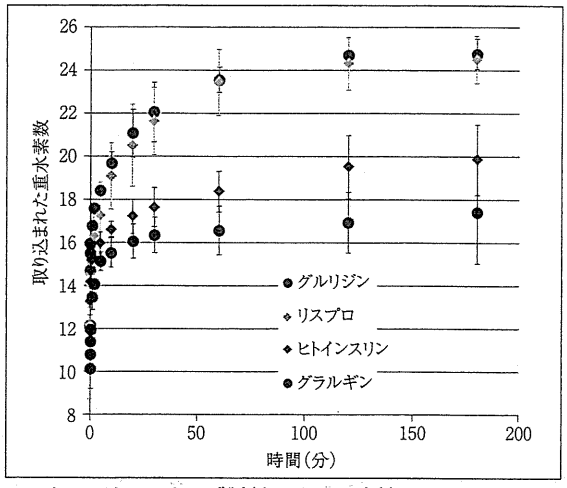


図5　インスリンアナログ製剤のHDX反応性
製剤原液を9倍容量の重水溶媒と混合して氷上でHDX反応後，反応液と同容量の1％ギ酸/50%アセトニトリルで反応を停止させた。
トラップカラム，C18 trap；溶媒，0.1%ギ酸，50%アセトニトリル；MS装置，LTQ-FT(Thermo Fisher Scientific)；流速，50μL/分

しやすくなるため，超速効型インスリンアナログは速やかに作用を発現するとされている。持続型インスリンアナログであるインスリンスペースグラルギン(グラルギン)は，A21のAsnがGly残基に置換され，さらにB鎖のC末端に2分子のArg残基が付加されており，これらの置換によりグラルギンのpI(6.7)は他のインスリンアナログ(5～5.5)よりも高い。中性付近のpIをもつグラルギンは投与後，中性pHの皮下組織で析出するため六量体から単量体への解離が遅くなり，作用を持続すると考えられている。このように多量体安定性はインスリンアナログ製剤の有効性に影響を及ぼす重要な物理的化学的性質の1つとなっており，簡便かつ迅速に多量体安定性の評価は重要である。

生体内のタンパク質は，柔軟性(揺らぎ)を保持した分子として存在しており，立体構造の揺らぎは多量体安定性に関係することが知られている。タンパク質の揺らぎ(動的構造)を解析する手法の1つとして，水素／重水素交換(HDX)反応とMSを組み合わせたHDX/MSがある[6,7](**図4**)。重水中でタンパク質のアミド水素は徐々に重水素に交換される。交換された水素数は，反応溶液のpHを酸性(pH 2.5)に変えてHDX反応を停止した後，MSで分析することで質量変化数として確認することができる。一般に分子表面の溶媒と接触しやすいアミド水素ほど速く重水素に交換され，水素結合に関与する水素原子や分子内部に存在するアミド水素などの交換速度は遅いことから，HDX/MSにより多量体安定性の異なるインスリンアナログの揺らぎを比較できる。

図5は，超速効型，速効型，および持続型インスリンアナログ製剤のHDX/MSを行い，重水素交換数を時間に対してプロットした結果である[8]。超速効型のリスプロおよびグルリジンはより多くの重水素を取り込み，持

表2 HDXパラメータ

	水素交換数			
	D_∞	D_s	D_m	D_f
ヒトインスリン	20.3	3.9	1.5	3.0
リスプロ	24.5	7.3	2.6	2.5
グルリジン	24.7	7.2	2.6	2.8
グラルギン	18.0	2.0	1.8	4.2

フィッティングに使用した式1 $D_t=D_\infty-D_f\exp(-k_ft)-D_m\exp(-k_mt)-D_s\exp(-k_st)$（$D_t$：反応開始から$t$ min後の交換数，t：反応時間（min），D_∞：最大交換数，$D_f/D_m/D_s$：各群由来の交換数，$k_f/k_m/k_s$：各群の水素の反応速度定数（min^{-1}））。各アナログの反応速度定数，$k_s\leq0.1$，$0.1<k_m\leq1$，$k_f>1$

D_∞，最大交換数；$D_f/D_m/D_s$，交換速度が遅い／中間／速い水素数

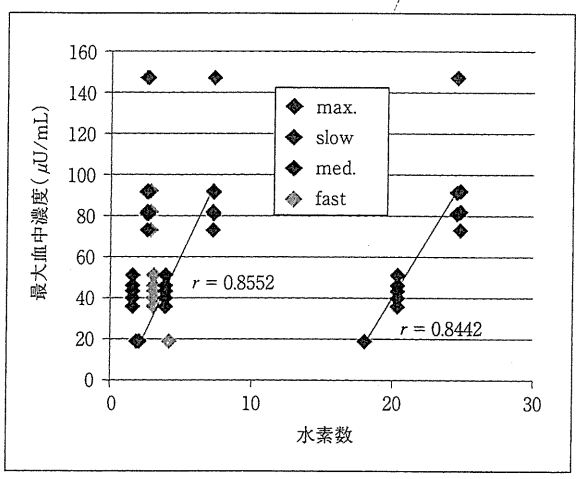


図6 HDXパラメータと最大血中濃度の相関性

Max.，D_∞；slow/med./fast.，$D_f/D_m/D_s$（表2参照）。最大血中濃度の値は，以下の論文を参照した。Mudaliar *et al.*（1999）*Diabetes Care* **22**（9）：1501-1506；Lindholm *et al.*（1999）*Diabetes Care* **22**（5）：801-805; Becker *et al.*（2008）*Clin. Pharmacokinet.* **47**（1）：7-20；EPAR "Product Information"（Apidra）；Product Monograph（Humalog），Eli Lilly Canada; Heise *et al.*（1998）*Diabetes Care* **21**（5）：800-803；Lepore *et al.*（2000）*Diabetes* **49**（12）：2142-2148.

効型のグラルギンと中間型製剤の重水素取り込み数は少ない。インスリンアナログ製剤のHDX反応性から，多量体安定性を確認することができる。また，HDX反応による重水素交換数の経時変化は，擬一次反応速度式で表すこと（フィッティング）が可能であり，その式から重水素最大交換数や反応速度の異なる重水素数などのHDXパラメータを求めることができる（**表2**）。製剤間のHDXパラメータの比較から，HDX反応性の差には最大交換数（D_∞）および交換速度の遅い水素数（D_s）を反映していることがわかる。さらに，HDXパラメータは最大血中濃度と相関があることが示唆された（**図6**）。

おわりに

本稿では，バイオ医薬品の物理的化学的性質解析の現状と最新技術について，実例を紹介しながら概説した。近年，生物活性，血中安定性，免疫原性および分子標的性の改善を目的とした改変・融合タンパク質，糖鎖改変タンパク質，PEG化ペプチド・タンパク質，細胞毒性が高い低分子化合物をカップリングさせた抗体医薬品など，新しい分子構造をもつ医薬品が次々と開発されている。これらの医薬品には，新規な構造および高度な不均一性をもち，ヒトでの薬理作用や安全性の予測が難しいものも少なくない。物理的化学的性質解析は，ファーストインマン試験の安全性確保のためにも重要性が増していくものと予想される。

■参考文献

1) 厚生労働省：第十六改正日本薬局方（http://www.pmda.go.jp/kyokuhou/3-24kokuji_index.htm.
2) Yanagihara S, Taniguchi Y, Hosono M, Yoshioka E, Ishikawa R, Shimada Y, Kadoya T, Kutsukake K.：Measurement of sialic acid content is insufficient to assess bioactivity of recombinant human erythropoietin., *Biol Pharm Bull*, **33**, 1596-1599（2010）
3) Maeda, E., Urakami, K., Shimura, K., Kinoshita, M. and Kakehi, K.：Charge heterogeneity of a therapeutic monoclonal antibody conjugated with a cytotoxic antitumor antibiotic, calicheamicin, *J Chromatogr A*, **1217**, 7164-7171（2010）
4) Moosmann, A., Christel, J., Boettinger, H. and Mueller, E.：Analytical and preparative separation of PEGylated lysozyme for the characterization of chromatography media, and *J Chromatogr A*, **1217**, 209-215（2010）
5) Ohno, A., Kawasaki, N., Fukuhara, K., Okuda, H. and Yamaguchi, T.：Time-dependent changes of oxytocin using（1）H-NMR coupled with multivariate analysis：A new approach for quality evaluation of protein/peptide biologic drugs, *Chem Pharm Bull*（*Tokyo*）, **57**, 1396-1399（2009）
6) Hamuro, Y. *et al.*：Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry, *J Biomol Tech*, **14**, 171-182（2003）
7) Chitta, R. K., Rempel, D. L., Grayson, M. A., Remsen, E. E. and Gross, M. L.：Application of SIMSTEX to oligomerization of insulin analogs and mutants, *J Am Soc Mass Spectrom*, **17**, 1526-1534（2006）
8) Nakazawa, S., Hashii, N., Harazono, A. and Kawasaki, N.：Analysis of oligomeric stability of insulin analogs using hydrogen/deuterium exchange mass spectrometry, *Anal Biochem.* **420**（1）, 61-67（2011）

AIDS & Clinical Research

Maeda, et al. J AIDS Clinic Res 2011, S2
http://dx.doi.org/10.4172/2155-6113.S2-004

Review Article Open Access

In vitro and *In vivo* Resistance to Human Immunodeficiency Virus Type 1 Entry Inhibitors

Yosuke Maeda[1], Kazuhisa Yoshimura[2], Fusako Miyamoto[3], Eiichi Kodama[3], Shigeyoshi Harada[2], Yuzhe Yuan[5], Shinji Harada[1] and Keisuke Yusa[5]*

[1]*Department of Medical Virology, Graduate School of Medical Virology, Kumamoto University, Honjo 2-1-1, Kumamoto 860-5886, Japan*
[2]*Center for AIDS Research, Kumamoto University, Kumamoto 860-0811, Japan*
[3]*Division of Emerging Infectious Diseases, Department of Internal Medicine, Tohoku University School of Medicine, Sendai, Japan*
[4]*Transfusion transmitted Diseases Center, Institute of Blood Transfusion, Chinese Academy of Medical Science, #26 Hua-Cai Road, Long-Tan-Si Industrial Park, Chenghua District, Chengdu, 610052 Sichuan Province, P. R. China*
[5]*Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, Kami-youga 1-18-1, Setagaya, Tokyo 158-8501, Japan*

Abstract

Viral entry is one of the most important targets for the efficient treatment of Human immunodeficiency virus type 1 (HIV-1)-infected patients. The entry process consists of multiple molecular steps: attachment of viral gp120 to CD4, interaction of gp120 with CCR5 or CXCR4 co-receptors, and gp41-mediated fusion of the viral and cellular membranes. Understanding the sequential steps of the entry process has enabled the production of various antiviral drugs to block each of these steps. Currently, the CCR5 inhibitor, maraviroc, and the fusion inhibitor, enfuvirtide, are clinically available. However, the emergence of HIV-1 strains resistant to entry inhibitors, as commonly observed for other classes of antiviral agents, is a serious problem. In this review, we describe a variety of entry inhibitors targeting different steps of viral entry and escape variants that are generated *in vitro* and *in vivo*.

Keywords: CD4-gp120 binding inhibitor; CCR5 antagonist; CXCR4 antagonist; Fusion inhibitor; Resistance; HIV-1

Introduction

The development of chemotherapy with antiretroviral agents has reduced the morbidity and mortality of Human immunodeficiency virus type 1 (HIV-1)-infected individuals. Successful treatment of HIV-1-infected patients using chemotherapy is partly due to a combination of different classes of antiviral agents against the viral protease or reverse transcriptase. However, successful eradication of the virus from infected individuals has not been achieved by antiviral treatment, and is often limited by the emergence of drug-resistant HIV-1 strains [1-3]. These problems highlight the need to develop novel anti-HIV-1 drugs that target different steps of the viral replication process. Viral entry is currently one of the most attractive targets for the development of new drugs to control HIV-1 infection. Viral entry proceeds through Env (gp120, gp41)-mediated membrane fusion, and consists of sequential steps: (i) attachment of viral gp120 to the CD4 receptor; (ii) binding of gp120 to CCR5 or CXCR4 co-receptors; and (iii) fusion of the viral and cellular membranes (Figure 1). A large number of inhibitors targeting different steps of the viral entry process have been developed, including peptides/peptide mimics, small molecules, and monoclonal antibodies (MAb).

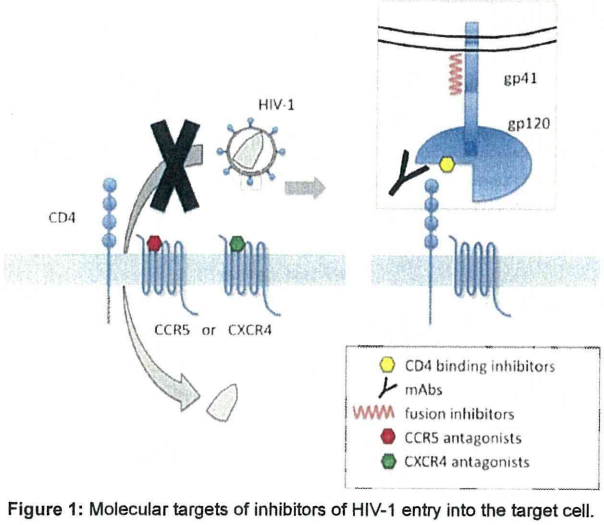


Figure 1: Molecular targets of inhibitors of HIV-1 entry into the target cell.

Enfuvirtide (also known as T-20) was the first of a new class of drugs known as fusion inhibitors, which was approved by the U.S. Food and Drug Administration (FDA) in 2003. Approval was given for the use of this drug in combination with other anti-HIV-1 medications to treat advanced HIV-1 infection in adults and children aged six years and older. The drug is an antiviral peptide that prevents HIV-1 entry by blocking gp41-mediated fusion [4-6]. Small compounds that can bind to the pockets of the extracellular loops of a coreceptor are expected to be potent antiviral agents. Several small-molecule CCR5 inhibitors have progressed through clinical development [7]. Maraviroc [8,9], a CCR5 antagonist, is the second entry inhibitor approved by the FDA in 2007 for treatment-experienced patients infected with a CCR5-tropic (R5-tropic) virus. Extensive research is currently underway to develop the next generation of entry inhibitors, however, the emergence of viral strains resistant to entry inhibitors, as well as other classes of antiviral

***Corresponding author:** Keisuke Yusa, Division of Biological Chemistry and Biologicals, National Institute of Health Sciences, Kami-youga 1-18-1, Setagaya, Tokyo 158-8501, Japan, Tel: +81-3-3700-1141 ext. 335; Fax: +81-3-3700-9084, E-mail: yusak@nihs.go.jp

Received October 18, 2011; **Accepted** December 02, 2011; **Published** December 05, 2011

Citation: Maeda R, Yoshimura K, Miyamoto F, Kodama E, Harada S, et al. (2011) *In vitro* and *In vivo* Resistance to Human Immunodeficiency Virus Type 1 Entry Inhibitors. J AIDS Clinic Res S2:004. doi:10.4172/2155-6113.S2-004.

agents, has been reported *in vitro* and *in vivo* [7,10]. In this review, we describe the current status of *in vitro* and *in vivo* resistance to HIV-1 entry inhibitors.

Resistance to CD4-gp120 binding inhibitors

Inhibition of CD4-gp120 binding: Entry of HIV-1 into target cells is mediated by the trimeric envelope glycoprotein complex, each monomer consisting of a gp120 exterior envelope glycoprotein and a gp41 transmembrane envelope glycoprotein [11]. Attachment of HIV-1 to the cell is initiated by the binding of gp120 to its primary CD4 receptor, which is expressed on the surface of the target cell. The gp120-CD4 interaction induces conformational changes in gp120 that facilitate binding to additional coreceptors (for example, CCR5 or CXCR4). Attachment inhibitors are a novel class of compounds that bind to gp120 and interfere with its interaction with CD4 [12]. Thus, these agents can prevent HIV-1 from attaching to the CD4+ T cell and block infection at the initial stage of the viral replication cycle (Figure 1). There are two primary types of HIV-1 attachment inhibitors: nonspecific attachment inhibitors and CD4-gp120 binding inhibitor [13].

In this section, we focus on the CD4-gp120 binding inhibitors, the soluble form of CD4 (sCD4), a fusion protein of CD4 with Ig (PRO542), a monoclonal anti-CD4 antibody (Ibalizumab, formerly TNX-355), CD4 binding site (CD4bs) monoclonal antibodies (b12 and VRC01), small-molecule HIV-1 attachment inhibitors (BMS-378806 and BMS-488043), and a new class of small-molecule CD4 mimics (NBD-556 and NBD-557) and a natural small bioactive molecule (Palmitic acid) (Figure 2). We also describe the resistance profiles against these CD4-gp120 binding inhibitors *in vivo* and/or in vitro.

Soluble CD4 (sCD4) and PRO542: In the late 1980s, various recombinant, soluble proteins derived from the N-terminal domains of CD4 were shown to be potent inhibitors of laboratory strains of HIV-1 [14]. Based on the potential of sCD4 to inhibit HIV-1 infection in vitro, this protein was tested for clinical efficacy in HIV-1-infected individuals; however, no effect on plasma viral load was observed [14]. Further examination revealed that doses of sCD4 significantly higher than those achieved in the clinical trial were required to neutralize primary clinical isolates of HIV-1, in contrast to the relatively sensitive, laboratory-adapted strains [15].

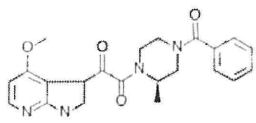
The first report of sCD4-resistant variants induced by *in vitro* selection showed that the resistant variant had a single mutation (M434T) in the C4 region [16]. During selection with sCD4, it was also reported that, seven mutations (E211G, P212L, V255E, N280K, S375N, G380R, and G431E) appeared during *in vitro* passage [17]. Further, a recombinant clone containing a V255E mutation was found to be highly resistant to sCD4 compared with the wild-type virus (114-fold higher 50% inhibitory concentration [IC_{50}] value). To determine the mutation profiles obtained during *in vitro* selection with sCD4, the atomic coordinates of the crystal structure of gp120 bound to sCD4 was retrieved from public protein structure database (PDB entry: 1RZJ). From these analyses, it was determined that almost all the described resistance mutations were located the inside the CD4-binding cavity of gp120 [17].

Recently, a novel recombinant antibody-like fusion protein (CD4-IgG2; PRO542) was developed in which the Fv portions of both the heavy and light chains of human IgG2 were replaced with the D1D2 domains of human CD4 [18]. PRO542 was shown to broadly and po-

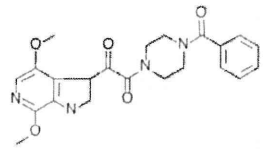
Profile of CD4-gp120 binding inhibitors

	Structure	Feature	Target	Resistant related mutations (region of gp160) [ref]
sCD4	Solbule form of CD4 domain1-4	First CD4-gp120 binding inhibitor	CD4 binding site of gp120	M434T (C4) [16], V255E(C2) [17]
PRO542	Tetravalent CD4 (domain1-2)-IgG	Developing for microbicide	CD4 binding site of gp120	* N/A
Ibalizumab	Anti-CD4 monoclonal antibody (MAb)	First-in-class, MAb inhibitor of CD4-mediated HIV entry	Domain 2 of CD4	N/A
b12	Anti-CD4 binding site Mab	Neutralizing around 40% of HIV-1 primary isolates	CD4 binding site of gp120	P369L (C3) [27]
VRC01	Anti-CD4 binding site Mab	Neutralizing over 90% of diverse HIV-1 primary isolates	CD4 binding site of gp120	K121A(C1), L179A(V2), T202A(C2), D279A(C2), R304A(V3), I420A(C4), I423A(C4), Y435A(C4), G471A (C5), D474A(C5) [31]
BMS-378806	see below Figure	First small molecule HIV-1 CD4 attachemnt inhibitor	CD4 binding site of gp120	V68A(C1), M426L(C4), M475I(V5), I595F(gp41) [33]
BMS-488043	see below Figure	improved in vitro antiviral activity and PK properties compared to BMS-378806	CD4 binding site of gp120	V68A(C1), L116I(C1), S375I/N(C3), M426L(C4) [34]
NBD-556	see below Figure	Inhibition of HIV-1 entry and enhancing neutralizing potency of Abs	CD4 binding site of gp120	S377N(C3), A433T(C4) [17], S375W(C3), I424A(C4), W427A(C4), V475A(C5) [38]
NBD-557	see below Figure	Inhibition of HIV-1 entry and enhancing neutralizing potency of Abs	CD4 binding site of gp120	N/A
Palmitic acid	$CH_3(CH_2)_{14}COOH$	A natural small bioactive molecule from *Sargassum fusiforme*	Domain 1 of CD4	N/A

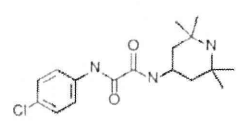
* N/A ; not available



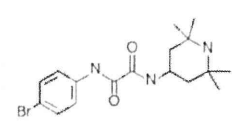
BMS-378806



BMS-488043



NBD-556



NBD-557

Figure 2: Profile of CD4-gp120 binding inhibitors including molecular structures of selected small molecular inhibitors.

tently neutralize HIV-1 subtype B isolates, and was also able to neutralize strains from non-B isolates with the same breadth and potency as for subtype B strains. PRO542 blocks attachment and entry of the virus into CD4+ target cells and were mainly developed for the prevention and transmission of HIV-1 through external application agents, such as microbicides.

Ibalizumab (TNX-355): Monoclonal anti-CD4 antibodies block the interaction between gp120 and CD4 and, therefore, inhibit viral entry [19]. Ibalizumab (formerly TNX-355) was a first-in-class, monoclonal antibody inhibitor of CD4-mediated HIV-1 entry [20]. By blocking CD4-dependent HIV-1 entry, ibalizumab was shown to be active against a broad spectrum of HIV-1 isolates, including recombinant subtypes, as well as both CCR5-tropic and CXCR4-tropic HIV-1 isolates. Many clinical trials with HIV-1-infected patients have demonstrated the antiviral activity, safety, and tolerability of ibalizumab. A nine-week phase Ib study investigating the addition of ibalizumab monotherapy to failing drug regimens showed transient reductions in HIV-1 viral loads and the evolution of HIV-1 variants with reduced susceptibility to ibalizumab. Further, clones with reduced susceptibility to ibalizumab contained fewer potential N-linked glycosylation sites (PNGSs) within the V5 region of gp120. Reduction in ibalizumab susceptibility due to the loss of V5 PNGSs was confirmed by site-directed mutagenesis [21].

Monoclonal antibodies, b12 and VRC01: Several broadly neutralizing MAbs isolated from HIV-1-infected individuals define conserved epitopes on the HIV-1 Env. These include the membrane proximal external region of gp41 targeted by MAbs 4E10 and 2F5 [22]; the carbohydrate-specific outer domain epitope targeted by 2G12 [23]; a V2-V3-associated epitope targeted by PG9/PG16 [24]; and the CD4bs [25] targeted by b12 and VRC01. The CD4bs overlaps with the conserved region on gp120 that is involved in the engagement of CD4. The prototypical CD4bs-directed MAb, b12, neutralizes around 40% of primary isolates, and its structure (in complex with the core of gp120) has been defined [26]. However, Mo et al. [27] reported the first resistant variant induced by *in vitro* selection with b12 that showed a P369L mutation in the C3 region of HIV-1_{JRCSF}. Further, several b12-resistant viruses commonly display an intact b12 epitope on the gp120 subunits [28], suggesting that quaternary packing of Env also confers resistance to b12.

A recently described CD4bs-directed MAb, VRC01, had been shown to be able to neutralize over 90% of diverse HIV-1 primary isolates [29]. The structure of VRC01 in complex with the gp120 core reveals that the VRC01 heavy chain binds to the gp120 CD4bs in a manner similar to that of CD4 [30]. The gp120 loop D and V5 regions contain substitutions uniquely affecting VRC01 binding, but not b12 or CD4-Ig binding. In contrast to the interaction of CD4 or b12 with the HIV-1 Env, occlusion of the VRC01 epitope by quaternary constraints was not a major factor limiting neutralization. Interestingly, many Ala substitutions at non-contact residues increased the potency of CD4- or b12-mediated neutralization; however, few of these substitutions enhanced VRC01-mediated neutralization [31]. This study suggests that VRC01 approaches its cognate epitope on the functional spike with less steric hindrance than b12 and, surprisingly, with less hindrance than the soluble form of CD4 itself. These differences might be related to the distinctly different angle of approach to the CD4bs employed by VRC01, in contrast to the more loop-proximal approach employed by CD4 and b12.

BMS-378806 and BMS-488043: BMS-378806 (Figure 2) is a recently identified small-molecule HIV-1 attachment inhibitor with good antiviral activity and pharmacokinetic properties [32]. BMS-378806 binds directly to gp120 with a stoichiometry of approximately 1:1 and with a binding affinity similar to that of soluble CD4. The potential BMS-378806 target site was localized to a specific region within the CD4 binding pocket of gp120 using HIV-1 gp120 variants carrying either compound-selected resistant substitutions or gp120-CD4 contact site mutations [32]. M426L (C4) and M475I (V5) substitutions located at or near gp120/CD4 contact sites were shown to confer high levels of resistance to the *in vitro* mutated HIV-1 variants, suggesting that the CD4 binding pocket of gp120 was the antiviral target. M434I and other secondary changes (V68A and I595F) also affect the drug susceptibility of recombinant viruses, presumably by influencing the gp120 conformation [33]. BMS-378806 (Figure 2) exhibited decreased, but still significant activity against subtype C viruses, low activity against viruses from subtypes A and D, and poor or no activity against subtypes E, F, G, and Group O viruses [33].

BMS-488043 (Figure 2) is a novel and unique small-molecule that inhibits the attachment of HIV-1 to $CD4^+$ lymphocytes. BMS-488043 exhibits potent antiviral activity against macrophage-, T-cell-, and dual-tropic HIV-1 laboratory strains (subtype B) and potent antiviral activity against a majority of subtype B and C clinical isolates [34]. Data from a limited number of clinical isolates showed that BMS-488043 exhibited a wide range of activity against the A, D, F, and G subtypes, with no activity observed against three subtype AE isolates [34]. The antiviral activity, pharmacokinetics, viral susceptibility, and safety of BMS-488043 were evaluated in an eight-day monotherapy trial that demonstrated significant reductions in viral load. To examine the effects of BMS-488043 monotherapy on HIV-1 sensitivity, phenotypic sensitivity assessment of baseline and post-dosing (day 8) samples were performed. The analyses revealed that four subjects showed emergent phenotypic resistance. Population sequencing and sequence determination of the cloned envelope genes revealed five gp120 mutations at four loci (V68A, L116I, S375I/N, and M426L) associated with BMS-488043 resistance; the most common (substitution at the 375 locus) located near the CD4 binding pocket [35].

NBD-556 and NBD-557: Targeting the functionally important and conserved CD4bs on HIV-1 gp120 represents an attractive potential approach to HIV-1 therapy or prophylaxis. Recently, a new class of small-molecule CD4 mimics was identified [36-38]. These compounds, which include the prototypic compound, NBD-556, and its derivatives, mimic the effects of CD4 by inducing the exposure of the co-receptor-binding site on gp120 [17,39]. NBD-556 and -557 (Figure 2) show potent cell fusion and virus-cell fusion inhibitory activity at low (micromolar) concentrations. A mechanistic study showed that both compounds target viral entry by inhibiting the binding of gp120 to its cellular receptor, CD4. A surface plasmon resonance study showed that these compounds bind to unliganded HIV-1 gp120, but not to CD4 [37]. Another recent study identified NBD-analogs as CD4 mimetics that were used for the prophylaxis and treatment of HIV-1 infection [39]. These compounds inhibited HIV-1 transmission by inhibiting the binding of the natural ligand, CD4, and prematurely triggering the envelope glycoprotein to undergo irreversible conformational changes. NBD-556 binds to the F43 cavity, which is formed by binding of gp120 to the CD4 receptor in a highly conserved manner [17,39].

Recently, our group reported that NBD-556 has potent neutralizing antibody-enhancing activity toward plasma antibodies that cannot access neutralizing epitopes hidden within the trimeric Env, such as gp120-CD4 induced epitope (CD4i) and anti-V3 antibodies [17]. Therefore, to investigate the binding site of NBD-556 on gp120, we in-

duced HIV-1 variants that were resistant to NBD-556 in vitro. Two amino acid substitutions (S375N in C3 and A433T in C4) were identified at passage 21 in the presence of 50 μM NBD-556. The profiles of the resistance mutations after selection with NBD-556 and sCD4 were very similar with regard to their three-dimensional positions.

Elucidation of the detailed molecular mechanisms governing the interaction between gp120 and NBD compounds will enable the optimization and evaluation of this strategy in more complex biological models of HIV-1 infection. Consequently, we will continue to synthesize NBD analogs and search for drugs with greater potency to change the tertiary structure of the envelope glycoproteins and reduce host cytotoxicity [40,41].

Palmitic acid : Previous studies with whole *Sargassum fusiforme* (*S. fusiforme*) extract and with the bioactive SP4-2 fraction demonstrated inhibition of HIV-1 infection in several primary and transformed cell lines [42]. Palmitic acid (PA), which was isolated from the SP4-2 bioactive fraction, specifically block productive X4 and R5-tropic HIV-1 infection [43]. PA occupies a novel hydrophobic cavity on the CD4 receptor that is constrained by amino acids F52-to-L70 [44], which encompass residues that have been previously identified as a region critical for gp120 binding. PA is mainly developed as microbicides [45].

Resistance to CCR5 antagonists

CCR5 antagonists: The binding of HIV-1 to CD4 molecules induces conformational change in gp120, resulting in the recognition of either CCR5 or CXCR4 as a coreceptor for HIV-1 (Figure 1). It has been shown that CCR5-utilizing HIV-1 (R5 virus) is associated with human-to-human transmission that predominate throughout the infection, while CXCR4-utilizing HIV-1 (X4 virus) emerges during the late stage of infection in approximately half of HIV-1-infected individuals and is associated with disease progression [46]. Most strikingly, it had been shown that homozygous individuals having a 32-bp deletion in the CCR5 coding region (CCR5Δ32) were found to be resistant to R5 HIV-1 and remained apparently healthy [47,48]. These findings suggested that CCR5 would be an attractive therapeutic target for treating HIV-1 infection, although it is a host factor. Several small molecule compounds have been developed and were found to bind CCR5 and inhibit R5 virus replication [49-53]. Molecular studies using CCR5 mutants indicated that these compounds bind to a cavity formed by transmembrane helices of CCR5, and thereby inducing the conformational change in an allosteric manner that is not recognized by gp120 of HIV-1 [54-58]. Among these, TAK-779 (Figure 3) was the first compound developed [49] that could inhibit not only HIV-1 infection, but also binding of RANTES (CCR5 ligand) to CCR5-expressing cells at nanomolar concentrations, but was terminated due to poor oral bioavailability. Maraviroc (MVC, UK427, 857) (Figure 3), however, has been approved and used in the clinic for the treatment of HIV-1 infection [8]. Another promising drug, vicriviroc (VCV, SCH-D, SCH-417690) (Figure 3), recently completed phase III trials but has not yet been approved [53].

Resistance to CCR5 antagonists: Although CCR5 antagonists target

Profile of CCR5 antagonist-resistant mutants

drug	virus used		resistant-related mutations		references
	virus name or *in vivo*	subtype	V3	Non-V3	
AD101	CC1/85	B	H305R, H308P, A316V, G321E	none	[60, 78]
TAK-779	JR-FL$_{V3lib}$	B	I304V, H305N, I306M, F312L, E317D	none	[63]
TAK-652	KK	unknown	ND[a]	ND	[59]
VVC	CC1/85	B	none	G516V, M518V, F519I (gp41)	[69, 84, 85]
VVC	RU570	G	K305R, R315Q, K319T	P437S (C4)	[64, 81]
VVC	S91	D	Q315E, R321G	E328K, G429R (C4)	[65]
VVC	*in vivo*	C	K305R, T307I, F316I, T318R, G319E	none	[67]
MVC	CC1/85	B	A316T, I323V	ND	[61]
MVC	JR-FL$_{V3lib}$	B	I304V, F312W, T314A, E317D, I318V	T199K, T275M (C2)	[62]
MVC	*in vivo*	B	P/T308H, T320H, I322V	D407G, Δ[b]N386 (V4)	[66]

[a]ND, not determined; [b]Δ, deletion

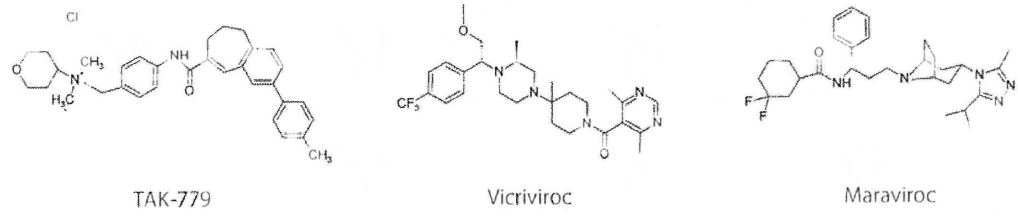


Figure 3: Profile of CCR5 antagonist-resistant mutants. The CCR5 antagonist-resistant mutants were isolated in vitro and in vivo across different subtypes of HIV-1. Resistance-related mutations were found in the V3 and non-V3 regions including the C2, V4, C4, and gp41. Chemical structures of representative CCR5 antagonists are shown.

a host cell receptor, the *in vitro* [59-64] and *in vivo* [65-67] emergence of viruses resistant to CCR5 antagonists in different subtypes has been reported, as shown in Figure 3. The most intuitive mechanism of resistance to CCR5 antagonists is likely to be the acquisition of CXCR4 use or selection of minority variants of CXCR4- or dual/mixed-tropic viruses [61,68-70]. Numerous studies showed that coreceptor selectivity of HIV-1 is primarily dependent on the third hypervariable region (V3 loop) of gp120 [71-74]. Furthermore, there is a simple rule to predict HIV-1 coreceptor usage called the 11/25 rule: if either the 11th or 25th amino acid position of V3 is positively charged, the virus will use CXCR4 as the coreceptor, otherwise it will use CCR5 [75]. Thus, a single amino acid substitution in the V3 loop is sufficient to acquire usage of CXCR4. However, these are rare cases when the viruses exclusively use CCR5.

Indeed, escape variants from selective pressure by natural ligand for CCR5, such as MIP-1α (CCL3) [76], or CCR5 antagonists [60], still use CCR5 and do not involve acquisition of CXCR4 usage. These studies indicate that acquisition of CXCR4 usage conferred by mutations in the V3 loop of gp120 results in the loss of replication fitness, as previously described [77]. However, the escape variants from CCR5 antagonists usually retain CCR5 usage [60,61,69,78], and recognize the antagonist-bound form of CCR5 as well as the free CCR5 form for entry by the accumulation of multiple amino acid mutations, called non-competitive resistance [61,79]. In non-competitive resistance, once saturating concentrations of antagonists were achieved, further inhibition was not observed, resulting in the plateau of inhibition, while competitive resistance can achieve inhibition of viral replication by a sufficient inhibitor concentration, resulting in a shift in the IC_{50} value (Figure 4). A principal determinant for the reduced sensitivity to CCR5 antagonists has been shown to be the V3 loop of gp120 although the mutations appear to be isolate-specific and antagonist-dependent [33].

In general, primary R5 viruses or laboratory-adapted R5 infectious clones cultured in stimulated peripheral mononuclear cells (PBMCs) have been used for the selection of CCR5 antagonist-resistant variants. However, the use of PBMCs for virus passage is donor-dependent and labor-intensive. Additionally, the use of a single clone for selection would need long-term passage to induce resistant viruses. To overcome these problems, we constructed R5-tropic infectious clones containing a V3 loop library, HIV-1_{V3Lib}. To construct replication competent HIV-1_{V3Lib}, we chose 10 amino acid positions in the V3 loop and incorporated random combinations of the amino acid substitutions derived from 31 subtype B R5 viruses into the V3 loop library (Figure 5). This novel *in vitro* system enabled the selection of escape variants from CCR5 antagonists over a relatively short time period.

In addition to the V3 library, we are currently using PM1/CCR5 cells for virus passages. The PM1/CCR5 cell line was generated by standard retrovirus-mediated transduction of parental PM cell line with the CCR5 gene, as previously described [63,76], and is highly sensitive to the R5 viruses compared to the parental PM1 cell line. Remarkably, the infection of PM1/CCR5 cells with R5 viruses induces prominent cell fusion, which is clear sign of virus proliferation. Thus, the use of PM1/CCR5 cells with the HIV-1_{V3Lib} allows us to focus on the contribution of the V3 loop in gp120 in CCR5 antagonist-resistance with a shortened selection period compared to the use of PBMCs with wild-type virus. As expected, we were able to isolate TAK-779- [63] and MVC-resistant [62] variants using replication competent HIV-1_{V3Lib}. Indeed, TAK-779- and MVC-resistant variants were determined to contain several amino acid substitutions within the V3 loop sequence. However, MVC-resistant variants also contained several amino acid substitutions in non-V3 regions (T199K and T275M), such as elsewhere in the gp120 to retain infectivity [80,81]. However, these mutations could not confer non-competitive resistance, indicating the importance of the V3 loop for non-competitive resistance.

Mechanisms of resistance: It is thought that docking of gp120 to CCR5 without CCR5 antagonists involves interactions of both the V3 tip with the second extracellular loop of CCR5 (ECL2) and the V3 stem-C4 region (bridging sheet) with the CCR5 N-terminus (NT) [82]. Since small molecule inhibitors interact with the pocket formed by transmembrane helices, thereby inducing allosteric conformational change in the ECL2, the wild-type virus can no longer interact with the ECL2. It is assumed that binding of small molecule inhibitors alters orientation between the ECL2 and NT regions, disrupting multipoint binding sites for gp120, thereby impeding gp120-CCR5 interaction [83]. Indeed, studies using CCR5 mutants showed that the escape variants were more dependent on tyrosine-sulfated CCR5 NT than wild-type viruses [65,66,84]. Furthermore, these escape variants were more sensitive to monoclonal antibodies recognizing the NT portion of CCR5 [65]. These studies indicated that the escape variants from CCR5 antagonists showed enhanced interactions with the NT that may be a consequence of a weakened interaction with the ECL2 (Figure 6).

Another genetic pathway is independent of V3 mutations. Vicriviroc-resistant mutants have been developed with multiple amino acid substitutions throughout the gp120 spanning the C2-V5 region without any changes in the V3 loop [69]. Recently, three amino acid changes in the fusion peptide domain of gp41 have been shown to be responsible for resistance although the effect of these mutations was

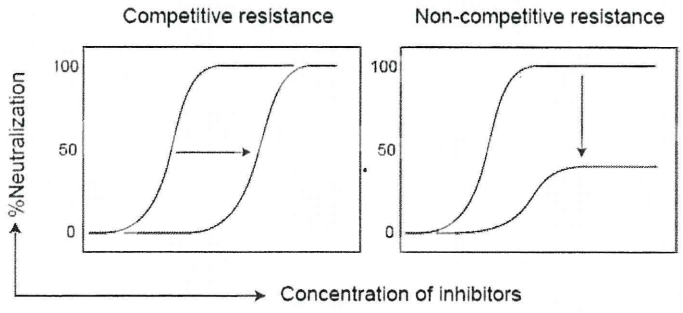


Figure 4: Typical competitive and non-competitive resistance profiles. Competitive resistance can achieve inhibition of viral replication by a sufficient inhibitor concentration, resulting in a shift in the IC50 value (left panel). In non-competitive inhibition, increasing concentrations of inhibitors have no effect, resulting in no increase in the inhibitory effect (right panel).

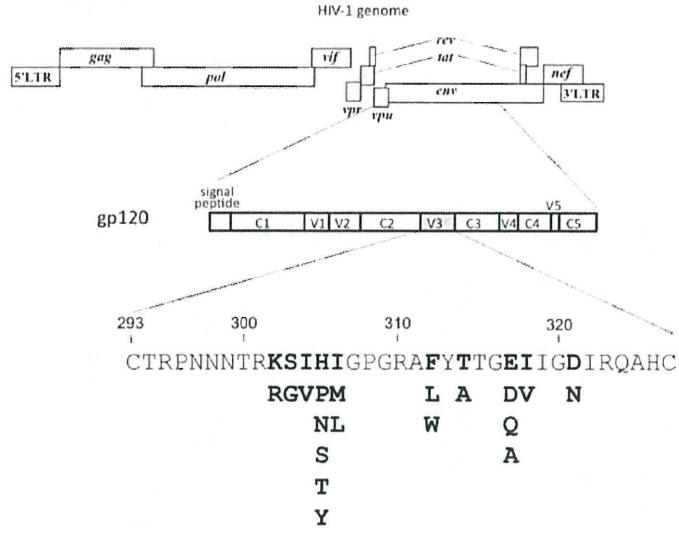


Figure 5: Schematic structure of HIV-1 V3 loop library showing introduced mutations in V3 for the analysis of escape mutants. Residues in boldface indicate the substitutions that were randomly incorporated in the V3 loop, possible >2 x 104 combinations. The amino acid substitutions were detected in 31 R5 clinical isolates.

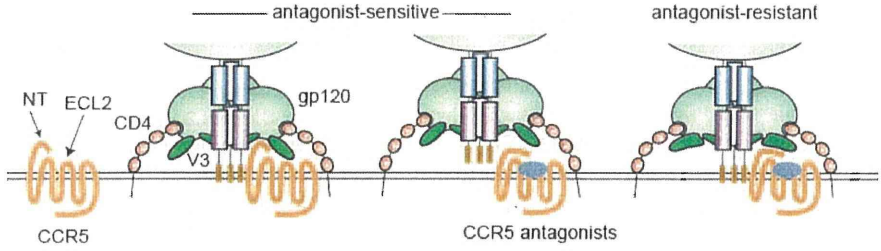


Figure 6: Resistant HIV-1 viruses can enter host cells in the presence of the CCR5 antagonist. The successful viral fusion requires the interaction of the V3 loop in gp120 with the ECL2 and NT of CCR5. CCR5 antagonists bind to the pocket formed by TM helices and induce allosteric conformational changes in the ECL2, thereby disrupting the interaction of gp120 with CCR5. The CCR5 antagonists-resistant viruses containing multiple amino acid substitutions in the V3 loop can recognize antagonist-bound forms of CCR5 by enhanced interaction with the NT.

context-dependent [84,85]. Thus, the mechanisms by which changes in the fusion peptide alter the gp120-CCR5 interaction still remain to be determined.

As previously mentioned, the patterns of mutations in escape variants against CCR5 antagonists were hypervariable and context-dependent, due in part to extensive sequence heterogeneity of HIV-1 *env*. Resistance to CCR5 antagonists was also found to be dependent upon cellular conditions such as cell tropism and the availability of CCR5. The differential staining of CCR5-expressing cells by various CCR5 monoclonal antibodies suggested that CCR5 exists in heterogeneous forms [86] and compositions of these multiple forms differed in cell type [87]. These findings suggested that different conformations of CCR5 with CCR5 antagonists might induce different substitutions in gp120. Moreover, the development of cross-resistance to other CCR5 antagonists is inconsistent, where some studies suggest that it may occur [69,78,79] and some suggest that it may not occur [61]. Additional data from *in vitro* and *in vivo* studies will be needed to elucidate the meaning of these studies.

Resistance to CXCR4 antagonists

CXCR4 as a target: CXCR4 is a coreceptor that is used for entry by X4-tropic viruses [88]; however, it is not always regarded as a suitable therapeutic target molecule for HIV-1 infection (Figure 1). R5 and X4 HIV-1 variants are both present in transmissible body fluids; however, R5-tropic HIV-1 transmits infection and dominates the early stages of HIV-1 pathogenesis [89], whereas X4-tropic HIV-1 evolves during the later stages and leads to acceleration of disease progression due to faster decline in $CD4^+$ T lymphocytes [90,91]. Coreceptor switching from CCR5 to CXCR4 occurs in approximately 40–50% of infected individuals [92]; in addition, the R5 virus is still present as a minor viral population even after emergence of the X4 virus. Furthermore, CXCR4 deletion in mice was shown to induce a variety of severe disorders and resulted in embryonic lethality [93], suggesting that CXCR4-targeting drugs may be less well tolerated than CCR5 inhibitors. These studies indicate that administration of CXCR4 inhibitors is relatively restricted to the later stage of infection after coreceptor switching. Therefore, the development of CXCR4 antagonists has proceeded at a deliberate pace when compared with that of other types of entry inhibitors.

Escape from CXCR4 antagonists: Based on the manner of escape of R5-tropic HIV-1 from CCR5 antagonists, four main resistance pathways may be intuitively possible for X4 HIV-1 escape from CXCR4 antagonists: (i) coreceptor switching from CXCR4 to CCR5; (ii) outgrowth of the pre-existing R5 virus; (iii) decrease in CXCR4 susceptibility by mutation(s) in Env; and (iv) utilization of the drug-bound